

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 15:40:56 ; Search time 8748.63 seconds
(without alignments)
3712.172 Million cell updates/sec

Title: US-09-715-036-4
Perfect score: 1853
Sequence: 1 atggctaagcttcttgagct.....aaaaaaaaaaaaaaaaaaaa 1853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 17159718 seqs, 8763200856 residues

Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

| | |
|-----|---|
| 1: | /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq.* |
| 2: | /cgn2_6/ptodata/2/pna/US06_COMB.seq.* |
| 3: | /cgn2_6/ptodata/2/pna/US07_COMB.seq.* |
| 4: | /cgn2_6/ptodata/2/pna/US080_COMB.seq.* |
| 5: | /cgn2_6/ptodata/2/pna/US081_COMB.seq.* |
| 6: | /cgn2_6/ptodata/2/pna/US082_COMB.seq.* |
| 7: | /cgn2_6/ptodata/2/pna/US083_COMB.seq.* |
| 8: | /cgn2_6/ptodata/2/pna/US084_COMB.seq.* |
| 9: | /cgn2_6/ptodata/2/pna/US085_COMB.seq.* |
| 10: | /cgn2_6/ptodata/2/pna/US086_COMB.seq.* |
| 11: | /cgn2_6/ptodata/2/pna/US087_COMB.seq.* |
| 12: | /cgn2_6/ptodata/2/pna/US088_COMB.seq.* |
| 13: | /cgn2_6/ptodata/2/pna/US089_COMB.seq.* |
| 14: | /cgn2_6/ptodata/2/pna/US090_COMB.seq.* |
| 15: | /cgn2_6/ptodata/2/pna/US091_COMB.seq.* |
| 16: | /cgn2_6/ptodata/2/pna/US092_COMB.seq.* |
| 17: | /cgn2_6/ptodata/2/pna/US093_COMB.seq.* |
| 18: | /cgn2_6/ptodata/2/pna/US094_COMB.seq.* |
| 19: | /cgn2_6/ptodata/2/pna/US095A_COMB.seq.* |
| 20: | /cgn2_6/ptodata/2/pna/US095B_COMB.seq.* |
| 21: | /cgn2_6/ptodata/2/pna/US095C_COMB.seq.* |
| 22: | /cgn2_6/ptodata/2/pna/US095D_COMB.seq.* |
| 23: | /cgn2_6/ptodata/2/pna/US096A_COMB.seq.* |
| 24: | /cgn2_6/ptodata/2/pna/US096B_COMB.seq.* |
| 25: | /cgn2_6/ptodata/2/pna/US096C_COMB.seq.* |
| 26: | /cgn2_6/ptodata/2/pna/US096D_COMB.seq.* |
| 27: | /cgn2_6/ptodata/2/pna/US096E_COMB.seq.* |
| 28: | /cgn2_6/ptodata/2/pna/US097A_COMB.seq.* |
| 29: | /cgn2_6/ptodata/2/pna/US097B_COMB.seq.* |
| 30: | /cgn2_6/ptodata/2/pna/US097C_COMB.seq.* |
| 31: | /cgn2_6/ptodata/2/pna/US098_COMB.seq.* |
| 32: | /cgn2_6/ptodata/2/pna/US099_COMB.seq.* |
| 33: | /cgn2_6/ptodata/2/pna/US6000_COMB.seq.* |
| 34: | /cgn2_6/ptodata/2/pna/US6001_COMB.seq.* |
| 35: | /cgn2_6/ptodata/2/pna/US6002_COMB.seq.* |
| 36: | /cgn2_6/ptodata/2/pna/US6003_COMB.seq.* |
| 37: | /cgn2_6/ptodata/2/pna/US6004_COMB.seq.* |
| 38: | /cgn2_6/ptodata/2/pna/US6005_COMB.seq.* |
| 39: | /cgn2_6/ptodata/2/pna/US6006_COMB.seq.* |
| 40: | /cgn2_6/ptodata/2/pna/US6007_COMB.seq.* |
| 41: | /cgn2_6/ptodata/2/pna/US6008_COMB.seq.* |
| 42: | /cgn2_6/ptodata/2/pna/US6009_COMB.seq.* |
| 43: | /cgn2_6/ptodata/2/pna/US6010_COMB.seq.* |

| | |
|-----|---|
| 44: | /cgn2_6/ptodata/2/pna/US6011_COMB.seq.* |
| 45: | /cgn2_6/ptodata/2/pna/US6012_COMB.seq.* |
| 46: | /cgn2_6/ptodata/2/pna/US6013_COMB.seq.* |
| 47: | /cgn2_6/ptodata/2/pna/US6014_COMB.seq.* |
| 48: | /cgn2_6/ptodata/2/pna/US6015_COMB.seq.* |
| 49: | /cgn2_6/ptodata/2/pna/US6016_COMB.seq.* |
| 50: | /cgn2_6/ptodata/2/pna/US6017_COMB.seq.* |
| 51: | /cgn2_6/ptodata/2/pna/US6018_COMB.seq.* |
| 52: | /cgn2_6/ptodata/2/pna/US6019_COMB.seq.* |
| 53: | /cgn2_6/ptodata/2/pna/US6020_COMB.seq.* |
| 54: | /cgn2_6/ptodata/2/pna/US6021_COMB.seq.* |
| 55: | /cgn2_6/ptodata/2/pna/US6022_COMB.seq.* |
| 56: | /cgn2_6/ptodata/2/pna/US6023_COMB.seq.* |
| 57: | /cgn2_6/ptodata/2/pna/US6024_COMB.seq.* |
| 58: | /cgn2_6/ptodata/2/pna/US6025_COMB.seq.* |
| 59: | /cgn2_6/ptodata/2/pna/US6026_COMB.seq.* |
| 60: | /cgn2_6/ptodata/2/pna/US6027_COMB.seq.* |
| 61: | /cgn2_6/ptodata/2/pna/US6028_COMB.seq.* |
| 62: | /cgn2_6/ptodata/2/pna/US6029_COMB.seq.* |
| 63: | /cgn2_6/ptodata/2/pna/US6030_COMB.seq.* |
| 64: | /cgn2_6/ptodata/2/pna/US6031_COMB.seq.* |
| 65: | /cgn2_6/ptodata/2/pna/US6032_COMB.seq.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|--------|---------------|--------|----|-----------------------|-------------------|
| 1 | 1853 | 100.0 | 1853 | 28 | US-09-715-036-4 | Sequence 4, Appli |
| 2 | 1391.2 | 75.1 | 1524 | 15 | US-09-141-220-5 | Sequence 5, Appli |
| 3 | 1391.2 | 75.1 | 1524 | 15 | US-09-141-220B-5 | Sequence 5, Appli |
| 4 | 1391.2 | 75.1 | 1524 | 15 | US-09-141-220C-5 | Sequence 5, Appli |
| 5 | 1391.2 | 75.1 | 1524 | 18 | US-09-494-096-6 | Sequence 6, Appli |
| 6 | 1377.4 | 74.3 | 1524 | 18 | US-09-455-294-7 | Sequence 7, Appli |
| 7 | 527.2 | 28.5 | 1446 | 13 | US-08-913-588-13 | Sequence 13, Appl |
| 8 | 527.2 | 28.5 | 1446 | 15 | US-09-108-010-13 | Sequence 13, Appl |
| 9 | 527.2 | 28.5 | 1446 | 29 | US-09-758-652-13 | Sequence 13, Appl |
| 10 | 516 | 27.8 | 1743 | 30 | US-09-762-381-1 | Sequence 1, Appli |
| 11 | 515.6 | 27.8 | 1488 | 13 | US-08-913-588-11 | Sequence 11, Appl |
| 12 | 515.6 | 27.8 | 1488 | 15 | US-09-108-010-11 | Sequence 11, Appl |
| 13 | 515.6 | 27.8 | 1488 | 29 | US-09-758-652-11 | Sequence 11, Appl |
| 14 | 510.8 | 27.6 | 1693 | 64 | US-60-312-544-1040 | Sequence 1040, Ap |
| 15 | 508.2 | 27.4 | 1458 | 13 | US-08-913-588-12 | Sequence 12, Appl |
| 16 | 508.2 | 27.4 | 1458 | 15 | US-09-108-010-12 | Sequence 12, Appl |
| 17 | 508.2 | 27.4 | 1458 | 29 | US-09-758-652-12 | Sequence 12, Appl |
| 18 | 355.8 | 19.2 | 1323 | 25 | US-09-654-617-258258 | Sequence 258258, |
| 19 | 355.8 | 19.2 | 1323 | 27 | US-09-684-016-258258 | Sequence 258258, |
| 20 | 352 | 19.0 | 1230 | 64 | US-60-312-544-1113 | Sequence 1113, Ap |
| 21 | 333.2 | 18.0 | 1801 | 25 | US-09-654-617-199 | Sequence 199, App |
| 22 | 333.2 | 18.0 | 1801 | 27 | US-09-684-016-199 | Sequence 199, App |
| 23 | 303 | 16.4 | 1830 | 25 | US-09-654-617-198 | Sequence 198, App |
| 24 | 303 | 16.4 | 1830 | 27 | US-09-684-016-198 | Sequence 189985, |
| 25 | 220.2 | 11.9 | 1855 | 25 | US-09-654-617-189985 | Sequence 189985, |
| 26 | 220.2 | 11.9 | 1855 | 27 | US-09-684-016-189985 | Sequence 189985, |
| 27 | 201.8 | 10.9 | 1551 | 13 | US-08-913-588-15 | Sequence 15, Appl |
| 28 | 201.8 | 10.9 | 1551 | 15 | US-09-108-010-15 | Sequence 15, Appl |
| 29 | 201.8 | 10.9 | 1551 | 29 | US-09-758-652-15 | Sequence 15, Appl |
| 30 | 201.6 | 10.9 | 3527 | 29 | US-09-759-967-1 | Sequence 1, Appli |
| 31 | 197.6 | 10.7 | 416 | 17 | US-09-304-517A-253748 | Sequence 253748, |
| 32 | 197.6 | 10.7 | 416 | 17 | US-09-371-146A-253748 | Sequence 253748, |
| 33 | 197.6 | 10.7 | 416 | 18 | US-09-440-687-15940 | Sequence 15940, A |
| 34 | 197.6 | 10.7 | 416 | 47 | US-60-144-084-4016 | Sequence 4016, Ap |
| 35 | 196 | 10.6 | 395 | 17 | US-09-304-517A-255598 | Sequence 255598, |
| 36 | 196 | 10.6 | 395 | 17 | US-09-371-146A-255586 | Sequence 255586, |
| 37 | 196 | 10.6 | 395 | 22 | US-09-565-240-2017 | Sequence 2017, Ap |
| 38 | 196 | 10.6 | 395 | 47 | US-60-141-135-1389 | Sequence 1389, Ap |
| 39 | 194.4 | 10.5 | 362 | 17 | US-09-304-517A-233545 | Sequence 233545, |
| 40 | 194.4 | 10.5 | 362 | 17 | US-09-371-146A-233545 | Sequence 233545, |
| 41 | 194.4 | 10.5 | 362 | 17 | US-09-391-630-1117 | Sequence 1117, Ap |

Db 1621 ttatccactaacataactttttgccacaaatgaataataataataagaataatgt 1680
QY 1681 agttttaatttttagtatgaataaagaataacaaaggggcattgatgcctttttgtttaaga 1740
Db 1681 agttttaatttttagtatgaataaagaataacaaaggggcattgatgcctttttgtttaaga 1740
QY 1741 tcggaatgtaacatatgtgcaatgagcagatatggagaaaaaccccttttcggggaaaaacat 1800
Db 1741 tcggaatgtaacatatgtgcaatgagcagatatggagaaaaaccccttttcggggaaaaacat 1800
QY 1801 gaataaaaaagaagttatgtctcacgcgcaaaaaaaaccccttttcggggaaaaacat 1853
Db 1801 gaataaaaaagaagttatgtctcacgcgcaaaaaaaaccccttttcggggaaaaacat 1853

RESULT 2
US-09-141-220-5
; Sequence 5, Application US/09141220A
; GENERAL INFORMATION:
; APPLICANT: Bannon, Gary A
; APPLICANT: Burks, A Wesley
; APPLICANT: Sampson, Hugh
; APPLICANT: Sosin, Howard
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: HS 102 PTO
; CURRENT APPLICATION NUMBER: US/09/141,220A
; CURRENT FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: PCT/US96/15222
; EARLIER FILING DATE: 1996-09-23
; EARLIER APPLICATION NUMBER: 60/074590
; EARLIER FILING DATE: 1998-02-13
; EARLIER APPLICATION NUMBER: 60/074624
; EARLIER FILING DATE: 1998-02-13
; EARLIER APPLICATION NUMBER: 60/074633
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Peanut
US-09-141-220-5

Art mls
not 100%

Query Match 75.1%; Score 1391.2; DB 15; Length 1524;
Best Local Similarity 95.4%; Pred. No. 1.8e-230;
Matches 1454; Conservative 0; Mismatches 68; Indels 2; Gaps 2;
QY 71 ggcagcagccggaggagaatgcgtgccaggttccagcgcctcaatgcgcagagacctgaca 130
Db 2 ggcagcaaccggaggagaacgcgtgccaggttccagcgcctcaatgcgcagagacctgaca 61
QY 131 accgcattgaatcggaggcggttacattgagaccttggaaccccccaacacccaggagttcg 190
Db 62 atcgcaattgaatcagaggcggttacattgagaccttggaaccccccaacacccaggagttcg 121
QY 191 aatgcgcggcggtgcgccctctctcgttagtctccgcgcgaacgcccttcgttaggcctt 250
Db 122 aatgcgcggcggtgcgccctctctcgttagtctccgcgcgaacgcccttcgttaggcctt 181
QY 251 tctactccaatgctccccagagagatcttcatccagcaagggaagggtatactttgggttga 310
Db 182 tctactccaatgctccccagagagatcttcatccagcaagggaagggtatactttgggttga 241
QY 311 tattecctggttgccttagcacatatgaagagcctgcacacaaaggacgcgcgatatacgt 370
Db 242 tattecctggttgccttagcacatatgaagagcctgcacacaaaggctcgtcgtatcagtc 301
QY 371 cccaaagaccaccaagacgttttgcaagaagaagaccaaagcccaacagcaacaagatagtc 430
Db 302 cccaaagaccaccaagacgtctctccaaaggagaagaccaaagcccaacagcaacaagatagtc 361

QY 431 accagaaggtgcaccggtttcaatgagggatgatctcattgcagttccccaccggtgtgtcctt 490
Db 362 accagaaggtgcaccggtttcgatgagggatgatctcattgcagttccccaccggtgtgtcctt 421
QY 491 tctggctgtacaacgaccagacactgatgttctgtgtgttcttcttactgacaccaaaca 550
Db 422 tctggctgtacaacgaccagacactgatgttctgtgtgttcttcttactgacaccaaaca 481
QY 551 acaacgacaacacagcttgatcagttccccaggagattcaatttggctgggaaccacgagc 610
Db 482 acaacgacaacacagcttgatcagttccccaggagattcaatttggctgggaaccacgagc 541
QY 611 aagagttcttaaggtaccagcaacaaagcagacaaagcagacgaagaagcttaccatata 670
Db 542 aagagttcttaaggtaccagcaacaaagcagacaaagcagacgaagaagcttaccatata 601
QY 671 gcccatagcccgcatagtcggccttagacgagagaagcgtgaatttcgcctcgaggac 730
Db 602 gcccatagcccgcaagtcagccttagacaaagaagcgtgaatttagcctcgaggac 661
QY 731 agcacagccgcagagagaagcagcagacaagaagaagacgaaggtggaaaaacatcttca 790
Db 662 agcacagccgcagagagaagcagcagacaagaagaagaagcgaaggtggaaaaacatcttca 721
QY 791 gcggcttcacgcgcggagttccttggaacaagccttccagggttgacgacagacagattgtgc 850
Db 722 gcggcttcacgcgcggagttccttggaacaagccttccagggttgacgacagacagattgtgc 781
QY 851 aaaaatctgtgggcgagaaacagagagtgaagaagagggagccattgtgacggtgaggggag 910
Db 782 aaaaacctaagaggcgagaccgagagagtgaagaagagggagccattgtgacagtgagggag 841
QY 911 gcctcagaatcttgagcccgagatggaacgagaggtgccgacgaagaagaggaatacagtg 970
Db 842 gcctcagaatcttgagcccgagatggaagaagcagcgtgccgacgaagaagaggaatacagtg 901
QY 971 aagatcaatatgaataccatgaacaggatggaagcgtggcagggggaagcagagcgggg 1030
Db 902 aagatgaatatgaatacagatgaagaggatagaagcgtggcagggggaagcagagcgagg 961
QY 1031 ggaatgggtattgaagagacgatctgcaccgcgatgtgttaaaaaagaacattggtagaaaaca 1090
Db 962 ggaatgggtattgaagagacgatctgcaccgcgaagtgtctaaaaagaacattggtagaaaaca 1021
QY 1091 gatccccctcacatctacgatccctcagcgtggttcaactcaaaactgcc-acgatctcaac 1149
Db 1022 gatccccctgacatctcaaacctca-acgtggttcaactcaaaactgccaaacgatctcaac 1080
QY 1150 cttctaatccttaggtggcttggaacttagtgcgtgaatatggaaatctctacaggaatgca 1209
Db 1081 cttctaatccttaggtggcttggaacttagtgcgtgaatatggaaatctctacaggaatgca 1140
QY 1210 ttgtttgtcctcactacaacacccacacagcatcatatataatggaatctctacaggaatgca 1269
Db 1141 ttgtttgtcctcactacaacacccacacagcatcatatataatggaatctctacaggaatgca 1200
QY 1270 gctcacgtgcaagtgggtggacagcaacgacacacagcagcagcagcagcagcagcagcagcag 1329
Db 1201 gctcacgtgcaagtgggtggacagcaacgacacacagcagcagcagcagcagcagcagcagcag 1260
QY 1330 ggtcacgttcttctgtgtgccacagaaacttcgcctggcgtgggaagtcacagcagcagcagcagcag 1389
Db 1261 ggtcacgttcttctgtgtgccacagaaacttcgcctgcgtgcgtgggaagtcacagcagcagcagcag 1320
QY 1390 ttcgaatacgtggcattcaagacagattcaaggccccagcagcagcagcagcagcagcagcagcagcag 1449
Db 1321 ttcgaatacgtggcattcaagacagactcaaggccccagcagcagcagcagcagcagcagcagcagcag 1380
QY 1450 aactccttcataagataacacctgcgcggagggaggtgtgtgcaaaatcattatggtccctccaaagg 1509
Db 1381 aactccttcataagataacacctgcgcggagggaggtgtgtgcaaaatcattatggtccctccaaagg 1440

; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Sosin, Howard B.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: HS 102 CIP (2)
; CURRENT APPLICATION NUMBER: US/09/494,096
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 09/141,220
; PRIOR FILING DATE: 1998-08-27
; PRIOR APPLICATION NUMBER: 09/240,557
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 09/241,101
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 09/248,673
; PRIOR FILING DATE: 1999-02-11
; PRIOR APPLICATION NUMBER: 60/073,283
; PRIOR FILING DATE: 1998-01-29
; PRIOR APPLICATION NUMBER: 60/074,590
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 60/074,624
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/248,674
; PRIOR FILING DATE: 1999-02-11
; PRIOR APPLICATION NUMBER: 60/122,566
; PRIOR FILING DATE: 1999-03-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-494-096-6

Query Match 75.1%; Score 1391.2; DB 18; Length 1524;
Best Local Similarity 95.4%; Pred. No. 1.8e-230;
Matches 1454; Conservative 0; Mismatches 68; Indels 2; Gaps 2;

QY 71 ggcagcagccggagagagaatgcgtgccagttccagcgccctcaatggcgagagacctgaca 130
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2 ggcagcaaccggagagagaacgcgtgccagttccagcgccctcaatggcgagagacctgaca 61

QY 131 accgcattgaatcgaggggcggttacattgagaccttggaaccccaacaaacacaggagttcg 190
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 62 atcgattgaatcagaggggcggttacattgagaccttggaaccccaacaaacacaggagttcg 121

QY 191 aatgcgcggcgctgcacctctctcgtcttagtctccgcgcaacgcaccttgtaggcctt 250
Db 122 aatgcgcggcgctgcacctctctcgtcttagtctccgcgcaacgcaccttgtaggcctt 181

QY 251 ttactccaatgctcccaggagatcttcatcagcaggaaggggatactttgggttg 310
Db 182 ttactccaatgctcccaggagatcttcatcagcaggaagggatactttgggttg 241

QY 311 tattccctggttgctcctagcacatatgaagagcctgcacaacaaggcgcgatatcagt 370
Db 242 tattccctggttgctcctagcacatatgaagagcctgcacaacaaggctcgtatctcagt 301

QY 371 cccaaagaccaccagacgctttgcaagaagaagaccacaaagcccaacagacagatagtc 430
Db 302 cccaaagaccaccagacgctctccaaggagagaccacaaagcccaacagacagatagtc 361

QY 431 accagaagggtgcaccgtttcaatgaggggtgatctcattgcagttccaccgggttgctt 490
Db 362 accagaagggtgcaccgtttcgatgaggggtgatctcattgcagttccaccgggttgctt 421

QY 491 tctggctgtacaacgacacgacactgatgtgtgtctgttctcttaactgacaccaaca 550
Db 422 tctggctgtacaacgacacgacactgatgtgtgtgtgttcttcttactgacaccaaca 481

QY 551 acaacgacaaccagcttgatcagttccccagagagattcaatttggctgggaaccacgagc 610
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

; Sequence 7, Application US/09455294
; GENERAL INFORMATION:
; APPLICANT: Bannon, Gary A.
; APPLICANT: Burks, Wesley A.
; APPLICANT: Caplan, Michael J.
; APPLICANT: Sampson, Hugh
; APPLICANT: Sosin, Howard
; TITLE OF INVENTION: Peptide Antigens
; FILE REFERENCE: 2002834-0004
; CURRENT APPLICATION NUMBER: US/09/455,294
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 7
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-455-294-7

Query Match      74.3%;   Score 1377.4;   DB 18;   Length 1524;
Best Local Similarity 94.9%;   Pred. No. 4.4e-228;
Matches 1445;   Conservative 0;   Mismatches 76;   Indels 2;   Gaps 2;

QY 72 gcagcagccggaggagaaatcggtgccagttccagcgcctcaatgcgcagagacacctgacaa 131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 gcagcaaccggagagaacgcgtgccagttccagcgcctcaatgcgcagagacacctgacaa 62

QY 132 ccgcattgaatcgcaggaggcggttacattgagacttggaaaccccaaacaccaggagtcca 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 tcgcattgaatcagagccggttacattgagacttggaaaccccaaacaccaggagtcca 122

QY 192 atgcgcggcgctgcgcctctctctcgcttagtccctccgcgcgaaccccttcgtagccctt 251
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 atgcgcggcgctgcgcctctctctcgcttagtccctccgcgcgaaccccttcgtagccctt 182

QY 252 ctactccaatgctccccaggagatcttccatccagcaagggaaggatactttgggttgat 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 ctactccaatgctccccaggagatcttccatccagcaagggaaggatactttgggttgat 242

QY 312 attccctggttgctccatagcacatatgaagagcctgcacaaacaggacgccgatcagtc 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 attcgctggttgctccatagacactatgaagagcctcacacaaagtcgctgatctcagtc 302

QY 372 ccaaagaccacaagacgtttgcaagaagaagaccaaagcccaacagcaacaagatagtc 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 ccaaagaccacaagacgtctccaaagagaagaccaaagcccaacagcaacaagatagtc 362

QY 432 ccagaaggtgcaccggttcaatgagggtgatctcattgcagttcccaccggtgtgcttt 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 ccagaaggtgcaccggttcgatgagggtgatctcattgcagttcccaccggtgtgcttt 422

QY 492 ctggctgtacaacgaccacgacactgatgttgctgtgtttctcttactgacacccaaca 551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 ctggctgtacaacgaccacgacactgatgttgctgtgtttctcttactgacacccaaca 482

QY 552 caacgacacaccagcttgatcagttccccaggagattccaatttggctgggaaccacgagca 611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 483 caacgacacaccagcttgatcagttccccaggagattccaatttggctgggaaccacgagca 542

QY 612 agagttcttaaggtaccagcaacaagcagacaaagcagacaagaagcttaccatatag 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 543 agagttcttaaggtaccagcaacaagcagacaaagcagacaagaagcttaccatatag 602

QY 672 ccatacagcccgcatagtcggccttagacgagaagcgtgaatttcgccctcaggac 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 603 ccatacagcccgcataagtcagccctagacaaagaagcgtgaatttagccctcaggaca 662

QY 732 gcacagccgcagagaacgagcagagacaagaagaagacgaaggtggaaacatcttcag 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 663 gcacagccgcagagaacgagcagagacaagaagaagaaacgaaggtggaaacatcttcag 722

QY 792 cggcttcacgccggagttccttgaaacaagccttccaggttgacgacagacagattgtgca 851
```

```

    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 723 cggcttcacgccggagttccttggaacaagctttccaggttgacgacagacagatagtcca 782

QY 852 aaatctgtggtggcgagaaacagagagtgaagaagaggagccattgtgacggtgaggggag 911
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 783 aaacctaagaggcgagactgagagtgaagaagaggagccattgtgacagtgaggggag 842

QY 912 cctcagaatcttgagcccagatggaaacgagaggtgcccagacgaagaaggaatacagatga 971
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 843 cctcagaatcttgagcccagatgaaagagacgtgcccagacgaagaaggaatacagatga 902

QY 972 agatcaatatgaataccatgaacaggatggaaggcgtggcagggggaagcagagcggggg 1031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 903 agatgaatatgaatacgtatgaagaggatagaacgcgtggcagggggaagcagagcgagg 962

QY 1032 gaatggtattgaagagacgatctgcaccgcctgtgttaaaaaaacaattggtggaaaacag 1091
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 963 gaatggtattgaagagacgatctgcaccgcctgtgttaaaaaaacaattggttagaaaacag 1022

QY 1092 atccctcacatctacgatcctcagcgctggttccactcaaaaactgcc-acgatctcaacc 1150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1023 atccctgacatctacaacccctca-agctggttccactcaaaaactgccaaacgatctcaacc 1081

QY 1151 ttctaactcttaggtggcttgacttagtgcagaatatggaaatctctacaggaatgc 1210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1082 ttctaactacttaggtggcttgaaacctagtgcagaatatggaaatctctacaggaatgc 1141

QY 1211 tgttgtccctcactacaacacacacacacacacacacacacacacacacacacacacac 1270
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1142 tgttgtcgtcactacaacacacacacacacacacacacacacacacacacacacacac 1201

QY 1271 ctacagtgcaagtggtggacagcaacgcgcacacagaggtgtacgacgaggttcaagagg 1330
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1202 ctacagtgcaagtcgtggacagcaacgcgcacacagaggtgtacgacgaggttcaagagg 1261

QY 1331 gtcacgttcttgtgtgcccacagaaacttcgcgcgtggctgggaaagtcacagagcgagaact 1390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1262 gtcacgtgcttgtgtgcccacagaaacttcgcgcgtggctgggaaagtcacagagcgagaact 1321

QY 1391 tcgaatacgtggcattcaagacagattcaaggcccagcatagccaactttgccggtgaaa 1450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1322 tcgaatacgtggcattcaagacagactcaaggcccagcatagccaacttcgccggtgaaa 1381

QY 1451 actccttcatagataaacctgccggagagggtggttgcaaatcctatggcctcccaagg 1510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1382 actccttcatagataaacctgccggagagggtggttgcaaatcctatggcctcccaagg 1441

QY 1511 agcaggcaaggcgagcttaagaacaacaaccccttcaagttcttcgttccaccttttcagc 1570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1442 agcaggcaaggcgagcttaagaacaacaaccccttcaagttcttcgttccaccttttcagc 1501

QY 1571 agtctccgagggtgtgtggttaa 1593
    ||||| ||||| ||||| ||||| |||||
Db 1502 agtctccgagggtgtgtggttaa 1524
```

```

RESULT 7
US-08-913-588-13
; Sequence 13, Application US/08913588
; GENERAL INFORMATION:
; APPLICANT: Fader, Gary M.
; APPLICANT: Kinney, Anthony J.
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; TITLE OF INVENTION: CLASSES OF SOYBEAN SEED
; TITLE OF INVENTION: PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
```


; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,588
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-913-588-13

Query Match 28.5%; Score 527.2; DB 13; Length 1446;
Best Local Similarity 62.2%; Pred. No. 2.9e-81;
Matches 983; Conservative 0; Mismatches 453; Indels 144; Gaps 4;

QY 13 cttgagctttcttttgcgttttgcgtttctagttctgtgggagctagcagcatctccttcagg 72
Db 10 CTGTGTTCTTTCCTTTGTTTCTGCTTTTCAGTGGTGCTGCTTCGCTTTCAGTTTCAGA 69
QY 73 cagcagccggagggaatcggtgccagttccagcgccctcaatgcgagagacctgacaac 132
Db 70 GAGCAGCCACAGCAAAACGAGTGCCAGATCCAACGCCCTCAATGCCCTAAACCCGGATAAC 129
QY 133 cgcattgaatcggaggcggttacattgagacttggaaaccccaacaacacaggagttcgaa 192
Db 130 CGTATAGAGTCAGAGGTGGCTTCATTGAGACATGGAACCCCTAACAAACAGCCATTCCAG 189
QY 193 tgcgcggcggtgcgcctctctcgtcttagtctccgcgcgcgaacgcccttcgttaggccttc 252
Db 190 TGTGCCGGTGTGGCCCTCTCTCGCTGCACCCCTCAACCGCAACGCCCTTCGCAGACCTCC 249
QY 253 tactccaatgctccccaggagatcttccatccagcaaggaaaggggatacttttgggttgata 312
Db 250 TACACCAACGCTCCCGAGGAGATCTACATCCAACAAGGTAGTGGTATTTTGGCATGATA 309
QY 313 ttccctgggtgtccttagcacatatgaagagccctgcaacaacaaggagcccgatatcagttcc 372
Db 310 TTCCCGGGTGTCTCTAGCACATTTGAAGAGCCTCAACAAAAGGA----- 354
QY 373 caaagaccaccaagacggtttgcaagaagaagaccaaagcccaagcaacaagatagtcac 432
Db 355 -----CAAAGCAGCAGGCCCTCAACAAACGCTCAC 381
QY 433 cagaaggtgcaccgtttcaatgaggtgatctcattgcaytctccaccgggtgttgcttcc 492
Db 382 CAGAAGATCTATCACTTCAGAGAGGGTGATTGATTGCAGTGCCAAACCGGTTTTCATAC 441
QY 493 tggctgtacaacgacacgacactgatgtgtgtgctgttcttcttactgacaccaacaac 552
Db 442 TGGATGTACAACAATGAAGACACTCCTGTGTGTGCCGTTTCTCTATTGACACCAACAGC 501
QY 553 aacgacaaccagcttgatcagttccccaggagattcaatttggctgggaacccacgagcaa 612
Db 502 TTCCAGAACCCAGCTCGACCATGATGCCCTAGGAGATTCTATCTTGCTGGGAACCAAGAGCAA 561

QY 613 gagttcttaaggtaccagcaacaagcagacaaagcagacgaagaagcttaccatatagc 672
Db 562 GAGTTTCTACAGTATCAGCCACAGAAGCAGCAA----- 594
QY 673 ccatacagccccgcatagtcggcctagacgagaagagcgtgaatttcgcctcgcaggacag 732
Db 595 -----GGAGGTACT 603
QY 733 cacagccgcagagaacgagcaggaacaagaagaagacgaaggtggaacacatcttcagc 792
Db 604 CAAAGCCAGAAAGGAAGCGTCAGCAAGAAGAAGAAACGAAGGAGGCAGCATATTGAGT 663
QY 793 ggcttcacgcccggagttccctggaaacaagcccttcagggttgacgacagacagattgtgcaa 852
Db 664 GGCTTCGCCCCCGGAATTCTTGGAACATCGGTTT---GTCGTGGACAGGAGATAGTGAGA 720
QY 853 aatctgtggggcgagaaacgagagtgaaagaagaggagccattgtgacggtgagggagggc 912
Db 721 AAGCTACAAGGTGAGAAACGAAGAGGAAGAGAGGGTGCCATTGTGACAGTGAAAGGAGGT 780
QY 913 ctcagaaatcttgagcccagatggaacgagaggtgcccagcgaagaagaagaaatacagatgaa 972
Db 781 CTCAGCGTGATAAGCCAC-----CCACGGGAAGAGCAGCAACAAGACCC 825
QY 973 gatcaatatgaataccatgaacagggatggaagggcgtggcaggggaagcagagggcggggg 1032
Db 826 GAGGAAGAGGAGAAAGCCAGATTGTGACGAGAAAGACAACATTGCCAAAGCCAAAGCAGA 885
QY 1033 aatggtattgaagagacgatctgcaccgcgatgtgttaaaaaagaacattggtggaacaaga 1092
Db 886 AATGGCATTGACGAGACCAATTGCACAATGAGACTTCGCCACAACATTGGCCACGACTTCA 945
QY 1093 tcccccaacatctacgatccctcagcgctggttcaactcaaaaactgccacgatctcaacctt 1152
Db 946 TCACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACAACCGCTACCAGCCTCGACTTC 1005
QY 1153 ctaatccttagtggtgcttagttagtctgtaatatggaatctctcacgaagaatgcattg 1212
Db 1006 CCAGCCCTCTCGTGGTCAAACTCAGTGCCCAAGTTGGATCACTCCGCAAGAATGCTATG 1065
QY 1213 ttgtccctcaactacaacaaccaacgcacacagcacaatataatgattgaggggagcgggct 1272
Db 1066 TTCGTGCCACACTACAACCTGAACCGCAACAGCATATAATATACGCATTGAATGGACGGGCA 1125
QY 1273 cactgcgaagtggtagacagcaacggcaacagagtgtagcagcagaggaggttcaagagggt 1332
Db 1126 TTGSTACAAGTGGTGAATTGCAATGGTGAGAGAGTGTTTGATGGAGAGCTGCAAGAGGGA 1185
QY 1333 cacttcttgggtgcccacagaacttcgccgtggtggggaagtccccagcagagcagaacttc 1392
Db 1186 CAGGTGTTAATTGTGCCACAAAACCTTTGCCGTGGCTGCAAGATCACAGAGCGACAACCTTC 1245
QY 1393 gaatacgtggcattcaagacagattcaagggcccaagcagcatagccaactttgccggtgaaaac 1452
Db 1246 GAGTATGTTTCATTCAAGACCAATGATAGACCCCTCGATCGGCAACCTTCAGAGGTGCAAAAC 1305
QY 1453 tccttcatagataaacctgcggagagggtggttgcaaaattcatatggcctcccaaggag 1512
Db 1306 TCATTGTTGAACGCATTGCCGGAGGAAGTGATTTCAGCAAACTTTTAACCTAAGGAGGCAG 1365
QY 1513 caggcaaggcagcttaagaacaacaaccccccttcaagttcttcgttccaccttttcagcag 1572
Db 1366 CAGGCCAGGCAGGTCAAGAAACAACAACCCCTTTTCAGCTTCTCTGGTTCCACCTTAAGGAGTCT 1425
QY 1573 tctccgagggcgtgtggctta 1592
Db 1426 CAGAGGAGAGTTGTGGCTTA 1445

RESULT 8
US-09-108-010-13
; Sequence 13, Application US/09108010
; GENERAL INFORMATION:

Db 1366 CAGGCCAGGAGGTCAGAACAAACCCCTTTCAGCTTCTCGTTCCACCTAAGGAGTCT 1425
QY 1573 tctccgagggctgtgctta 1592
Db 1426 CAGAGGAGAGTTGTGGCTTA 1445
RESULT 9
US-09-758-652-13
; Sequence 13, Application US/09758652
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-758-652-13
Query Match 28.5%; Score 527.2; DB 29; Length 1446;
Best Local Similarity 62.2%; Pred. No. 2.9e-81;
Matches 983; Conservative 0; Mismatches 453; Indels 144; Gaps 4;
QY 13 cttgagcttcttttgccttttgccttttgccttttgccttttgccttttgcctttcagg 72
Db 10 CTTGTTCTTCCCTTTGTTTCTGCTTTTTCAGTGGCTGCTGCTTTCGCTTTTCAGTTTCAGA 69
QY 73 cagcagccgagaggagaatgcgtgcccaggttccagcgcctcaatgcgcagagacctgacaac 132
Db 70 GAGCAGCCACAGCAAAACGAGTGCCAGATCCCAACGCCTCAATGCCCTAAACCCGGATAAC 129
QY 133 cgattgaatcggagggcggttacattgagacttggaaccccccaacacagggaggttcgaa 192
Db 130 CGTATAGAGTCAGAAAGGTGCTTCATTGAGACATGGAACCCCTAACAACAGCCATTCCAG 189
QY 193 tgcgcccgcgtcgccctctctcgttagtctcgcgcgcaacgccccttcgtaggccttc 252

Db 190 TGTGCCGGTGTGTGCCCTCTCTCGCTGCACCCCTCAACCGCAACGCCCTTCGCAGACCTTCC 249
QY 253 tactccaatgctccccaggagatcttcatccagcagaaggaaggatactttgggttgata 312
Db 250 TACACCAACGCTCCCCAGGAGATCTACATCCAACAAGGTAGTGGTATTTTGGCATGATA 309
QY 313 ttccctggttgcacatatgaagagccctgcacaacaaggagccgatatcagtc 372
Db 310 TTCCCGGGTTGTCTAGCACATTTGAAGAGCCTCAACAAAAAGGA----- 354
QY 373 caaagaccaccaagacgtttgcaagaagaagaccacaagccaacagcaacaagatagtcac 432
Db 355 -----CAAAGCAGCAGGCCCCCAAGACCGTCA 381
QY 433 cagaaggtgcaccgttttcaatgaggggtgatctcattgcagttccccaccggtgtgtcttc 492
Db 382 CAGAGATCTATCACITTCAGAGAGGGTGATTTGATTGCAGTGCCAACCCGTTTGCATAC 441
QY 493 tggctgtacaacgacaccacactgatgttgttgcctgttcttactgacaccaacaac 552
Db 442 TGGATGTACAACAATGAAGACACTCCTGTTGTCCTTCTTATTGACACCAACAGC 501
QY 553 aacgacaaccagcttgatcagttccccaggagattcaatttggctgggaaccacagagcaa 612
Db 502 TTCCAGAACCAGCTCGACCAGATGCTTAGGAGATTCTATCTTGTCTGGGAACCAAGAGCAA 561
QY 613 gatttotaaggtaccagcaacaagcagacaaagcagacgaagaagcgttaccatatagc 672
Db 562 GAGTTTCTACAGTATCAGCCACAGAGCAGCAA----- 594
QY 673 ccatacagcccgcatagtcgcctagacgagaagagcgtgaatttcgcccctcgaggacag 732
Db 595 -----GGAGGTACT 603
QY 733 cacagccgcagagaacgagcaggacaagaagaagacgaaggtggaaacatcttcagc 792
Db 604 CAAAGCCAGAAAGGAAAGCGTCAGCAAGAAGAAGAAACGAAGGAGGCAGCATATTGAGT 663
QY 793 ggcctcacgcccggagttcctggaacaagccttccaggttgacgacagacagattgtgcaa 852
Db 664 GGCTTCGCCCGGAATTCTTTGGAACATCGGTTCTC--GTCGTGGACAGGCAGATAGTGAGA 720
QY 853 aatctgtgggpcgagaaacgagagtgaaagaagggagccattgtgacggtgaggggagggc 912
Db 721 AAGCTACAAGGTGAGAACGAGAGGAAGAGAGGGTGCCATTGTGACAGTGAAAGGAGGT 780
QY 913 ctcaagaatcttgagcccagatggaacgagaggtgcccgaagaagaagaaatcacgatgaa 972
Db 781 CTCAGCGTGATAAGCCAC-----CCACGGAAGAGCAGCAACAAAGACCC 825
QY 973 gatcaatatgaataccatgaacaggatggaaggcgtggcaggggaagcagagggggggg 1032
Db 826 GAGGAAGAGGAGAGAGCCAGATTGTGACGAGAGAAAGACAAACATTCGCAAGCCAAAGCAGA 885
QY 1033 aatggtattgaagagacgatctgcaccgcgcatgtgttaaaaaaacattggtggaaacaga 1092
Db 886 AATGGCATTGACGAGACCCATTTGCACAATGAGACTTCGCCACAACATTTGCCCAGACTTCA 945
QY 1093 tccctcacatctacgatcctcagcgctggttccactcaaaactgccacgatctcaacctt 1152
Db 946 TCACCTGACATCTTCAACCCCTCAAGTGGTAGCATCAACAACCCGTACCAGCCCTCGACTTC 1005
QY 1153 ctaatccttaggtggttggacttagtgcgtaataatggaaaatctctacaggaatgcattg 1212
Db 1006 CCAGCCCTCTCGTGGTCAAACTCAGTGCCCGAGTTTGGATCTACTCCGCAAGAAATGCTATG 1065
QY 1213 ttgtccctcactacaacaccacgcacacagcatcatatggaatctctacagggaggggct 1272
Db 1066 TTGCTGCCACACTACAACCTGAACCGCAACAGCATAATATACGCATTGTAATGGACGGCA 1125
QY 1273 cagtgcaagtgggtggacagcaacggcaacagagtgtagcagcagagagcttcaaggggt 1332
Db 1126 TTGGTACAAGTGGTGAATTGCAATGGTGAGAGACTGTTTGATGGAGAGCTGCAAGAGGGA 1185


```

Db 1313 ctgcaagatcacagagtgacaacttcgagtagtgatgtgtcatcattcaagaccaatgatacaccca 1372
QY 1427 gcatagccaactttgcggtgaaactccttcattagataaacctgcgaggaggtggttg 1486
Db 1373 tgatcggcactcttcagggggcaactcattgttgaaacgcattaccagaggaagtgttc 1432
QY 1487 caaattcatatggcctcccaaggagcaggcaaggcagcttaagaaacaaaccccttca 1546
Db 1433 agcacactttcaacctaaagccagcagcaggccaggcagagataaagaaacaaaccccttca 1492
QY 1547 agttcttcgttcacacttttcagcagtgctccgagggtgtgtggttaaa 1594
Db 1493 agttctggttcacctcaggagtgctcagaagagagctgtgtggttaga 1540

RESULT 11
US-08-913-588-11
; Sequence 11, Application US/08913588
; GENERAL INFORMATION:
; APPLICANT: Fader, Gary M.
; APPLICANT: Kinney, Anthony J.
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; TITLE OF INVENTION: CLASSES OF SOYBEAN SEED
; TITLE OF INVENTION: PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: USA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913,588
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-4931
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-913-588-11

Query Match 27.8%; Score 515.6; DB 13; Length 1488;
Best Local Similarity 62.3%; Pred. No. 3e-79;
Matches 1001; Conservative 0; Mismatches 449; Indels 156; Gaps 6;

QY 17 agctttcttttgcgtttcttagttctctggtgagctagcagcatctcctt-----ca 70
Db 8 AGCTAGTTTTTCCCTTTGTTTCTGCTTTTCAGTGGCTGCTGCTTCGCTTTCAGTTCCA 67

QY 71 ggcagcagccggagagaatgcgtgccaggttcagcgcctcaatgcgcagagacctgaca 130
Db 68 GAGAGCAGCCTCAGCAAAACGAGTGCCAGATCCAAAAACTCAATGCCCTCAAAACCGGATA 127
```

```

QY 131 accgattgaatcggaggcggttacattgagacttggaaccccccaacacagaggttcg 190
Db 128 ACCGTATAGAGTCAGAAAGGAGGGCTCATTGAGACATGGAACCCCTAACAAAGCCATTCC 187
QY 191 aatgcgcggcgctgcgccctctctcgcttagtcctccgcgcgcaacgcccttcgtagccctt 250
Db 188 AGTGTGCCGGTGTGGCCCTCTCTCGCTGCACCCCTCAACCGCAACGCCCTTCGTAGACCTT 247
QY 251 tctaactccaatgctccccagagagatcttcataccagaaaggaagggatactttgggttga 310
Db 248 CTTACACCAACGGTCCCAGGAAATCTACATCCAACAAGGTAAAGGTATTTTGGCATGA 307
QY 311 tattccctggttgcctcagacatatgaagagcctgcacaacaaaggacgcogatatcagt 370
Db 308 TATACCCGGGTTGTCTAGCACATTTGAAGAGCCCTCAACAAC----- 349
QY 371 cccaaagaccaccaagacgttttgcaagaagaagaccaaagccaaacagcaacaagatagtc 430
Db 350 -----CTCAACAAAGAGGACAAAGCAGACAGACCACAAAGACCGGTC 388
QY 431 accagaaggtgcaccgtttcaatgaggtgatctcattgcagttccccaccggtgtgctt 490
Db 389 ACCAGAAGATCTATAACTTCAGAGAGGGTGATTTGATCGCAGTGCCCTACTGCTGTGTCAT 448
QY 491 tctggtgtacaacgaccacgacactgatgttggtgctgttcttcttactgacaccaaca 550
Db 449 GGTGGATGTACAACAATGAAGACACTCCTGTTGTTGTCGGCTTCTATTATTGACACCAACA 508
QY 551 aacacgacaaccagcttgatcagttccccagggagagattcaatttggctgggaaccacagac 610
Db 509 GCTTGGAGAACCAAGCTCGACCATGCCTAGGAGATCTATCTTGTGGGAACCAAGAGC 568
QY 611 aagagttcttaaggtaccagcaacaagcagacaaagcagacgaagaagccttaccatata 670
Db 569 AAGAGTTCTAAAATATCAGCAAGA----- 593
QY 671 gcccatacagcccgcatagtcggccttagacgagaagagcgtgaatttcgcctcctcaggagc 730
Db 594 -----GCAAGGAGGTC 604
QY 731 agcacagccgcagagaacgagcaggacaagaagaagacgaaggtggaaacatcttca 790
Db 605 ATCAAAGCCAGAAAGGAAGCATCAGCAAGAAGAAGAAAACGAAGGAGGCGCATATTGA 664
QY 791 gcggcttcacgcgcggagttcctggaaacaagccttcagggttgacgacagacagattgtgc 850
Db 665 GTGGCTTCACCCCTGGAATTTCTTGGAAACATGCATTCAGCGT---GGACAAGCAGATAGCGA 721
QY 851 aaaatctgtgggcgagaacgagagtgaaagaagaggagccattgtgacggtgaggggag 910
Db 722 AAAACCTACAAGGAGAGAAACGAAGGGGAAGACAAAGGAGGCCATTGTGACAGTGAAGGAG 781
QY 911 gcctcagaatcttgagccca-----gatggaacgagaggtgcgcgacgaag 955
Db 782 GTCTGAGCGTGATAAAACCCACCACGGACGACGACGACAAAGACCCCGAAGAGGAAG 841
QY 956 aagaggaatacगतgaagatcaata-----tgaataccatgaacaggatggaagcc 1006
Db 842 AAGAAGAAGAGGATGAGAAGCCACAGTGCAAGGGTAAAGACAAACACTGCCAACGCCCC 901
QY 1007 gtggcagggggaagcagagggcgggggggaatgggtattgaagagacgatctgcaccgcatgtg 1066
Db 902 GAGGAAGCCAAAGCAAAAGCAGAAAGAAATGGCATTTGACGAGACCATATGCAACCATGAGAC 961
QY 1067 ttaaaagaacattgtgtggaacagatccccctcacatctacgatccctcagcgtggttca 1126
Db 962 TTCGCCACAACATTGGGCCAGACTTCATCACCTGACATCTACAACCCCTCAAGCCCGTAGCG 1021
QY 1127 ctcaaaactgccacgatctcaaccttctaataccttagtggtggacttagtgctgaat 1186
Db 1022 TCACAACCGCCACCAAGCCCTTGACTTCCAGGCCCTCTCGTGGCTCAGACTCAGTGTGAGT 1081
```


| | | | |
|----|------|---|------|
| Qy | 1007 | gtggcaggggaagcagagcggggggaaatggattattgaagagacgatctgcaccgcatgtg | 1066 |
| Db | 902 | GAGGAAGCCAAAGCAAAAAGCAGAAGAAATGGCATTGACGAGACCATATGCACCATGAGAC | 961 |
| Qy | 1067 | ttaaaaaagaacatgggtggaacacagatccccctcacatctacgatcctcagcgctggttca | 1126 |
| Db | 962 | TTGCGCCACAACATTGGCCAGACTTCATCACCTTGACATCTACAACCCTCAAGCCGGTAGCG | 1021 |
| Qy | 1127 | ctcaaaaactgccaagatctcaaccccttctaactccttaggtggcttggacttagtgcgtgaat | 1186 |
| Db | 1022 | TCACAACCGCCACCAGCCTTGACTTCCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGT | 1081 |
| Qy | 1187 | atggaatatctacaggaatgcattgtttgtccctcactacaacacccaacacacacagca | 1246 |
| Db | 1082 | TTGGATCTCTCGCAAGAATGCAATGTTCTGTGCCACACTACAACCTGAACGCGAACAGCA | 1141 |
| Qy | 1247 | tcataatgcattgaggggacgggctcacgtgcaagtggtgacagcaacgccaacagag | 1306 |
| Db | 1142 | TAATATACGCATTGATGGACGGGCATTGATACAAAGTGGTGAATTGCAACGSGTGAGAGAG | 1201 |
| Qy | 1307 | tgtacgacgagggagcttcaagaggggtcacgttcttgtggtgccacagaaacttcgcggtgg | 1366 |
| Db | 1202 | TGTTTGATGGAGAGCTGCAAGAGGGACGGGTGCTGATCGTGGCCACAAAACCTTTGTGGTGG | 1261 |
| Qy | 1367 | ctgggaagtcccagagcgagaacttcgaaacgtgggcattcaagacagattcaaggccca | 1426 |
| Db | 1262 | CTGCAAGATCACAGAGTGACAACCTTCGAGTATGTGTCATTCAGAGACCAATGATACACCCA | 1321 |
| Qy | 1427 | gcatagccaaactttgcgggtgaaaaactccttcataagataaacctgccggaggagtggttg | 1486 |
| Db | 1322 | TGATCGGCACCTTGACGGGGCAAACTCATTTGTTGAACGCATTACCGAGAGGAAGTGATTTC | 1381 |
| Qy | 1487 | caaatccatataggcctcccaaggaggcaggcagcttaagaacaacaaccccccttca | 1546 |
| Db | 1382 | AGCACACTTTCAACCTTAAAAAGCCAGCAGGCCAGGCCAGAGATAAAGAACAACAACCCCTTCA | 1441 |
| Qy | 1547 | agttcttcgtccacaccttttcagcagtcctccgagggctgtggctta | 1592 |
| Db | 1442 | AGTTCCTGGTTCCACCTTCAGGAGTCTCAGAAGAGAGCTGTGGCTTA | 1487 |

RESULT 13

US-09-758-652-11
; Sequence 11, Application US/09758652
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971

Db. 665 GTGGCTTCACCCCTGGAAATCTTGGAACATGCATTAGCGT---GGACAAGCAGATAGCGA 721
QY 851 aaaaatctgtgggcgagaacgagagtgaagaagagggagccattgtgacgggtgaggggag 910
Db 722 AAAACCTACAAGGAGAGAACGAAGGGGAAGACAAGGGAGCCATTGTGACAGTGAAGGAG 781
QY 911 gcctcagaatcttgagccca-----gatggaacgagaggtgccgacgaag 955
Db 782 GTCTGAGCGTGTATAAACACCCACGGACGAGCAGCAACAAAGACCCCGAAGGAAG 841
QY 956 aagaggaatacatgaaagatcaata-----tgaataccatgaacaggatggaaggc 1006
Db 842 AAGAAGAAGAGGATGAGAAGCCACAGTGCAGGGTAAAGACAAACACTGCCAACGCCCC 901
QY 1007 gtggcaggggaagcagagagcggggggaatggtattgaagagacgatctgcaccgcagtgtg 1066
Db 902 GAGGAAGCCAAAGCAAAAGCAGAGAAATGGCATTGACGAGACCATATGCACCATGAGAC 961
QY 1067 ttaaaaaaacaattggttgaaacagatccocctcacatctacgatacctcagcgctgttca 1126
Db 962 TTCGCCACAACATTGGCCAGACTTTCATCACCTGACATCTACAAACCTCAAGCCGGTAGCG 1021
QY 1127 ctcaaaactgcacgacgatctcaaccttctaactccttaggtggttgacctagtgtgaat 1186
Db 1022 TCACAACCGCCACGACGCTTGACTTCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGT 1081
QY 1187 atggaatatcttacaggaatgcattgtttgtccctcactacaacaccaacgcacacagca 1246
Db 1082 TTGGATCTCTCGCAAGAATGCAATGTGCTGCCACACTACAACTGAACGGGAACAGCA 1141
QY 1247 tcatatatgcattgaggggacgggctcacgtgcaagtgtgagcagcaacggcaacagag 1306
Db 1142 TAATATACGGCATGAATGGACGGGCAATTGATACAAAGTGGTGAATTGCAACGGTGAGAG 1201
QY 1307 tgtacgacgaggagcttcaagaggggtcacgttcttgtgtggtgcoacagaacttcgccgtgg 1366
Db 1202 TGTTTGATGGAGAGCTGCAAGAGGGACGGGTGCTGATCGTGCACAAAACTTTGTGTFGG 1261
QY 1367 ctgggaagtccagagcgagaacttcgaatacgtggcattcaagacagattcaaggccca 1426
Db 1262 CTGCAAGATCACAGAGTGACAACTTCGAGTATGTGATTCAAGACCAATGATACACCCA 1321
QY 1427 gcatagccaaactttgccggtgaaaactccttcatagataaacctgccggaggtgtgtg 1486
Db 1322 TGATCGGCACTCTGTCAGGGGCAAACTCATTTGTAACGCATTACCAGAGGAAGTGATTC 1381
QY 1487 caaattcatatggcctcccaaggggagcaggaaggcagcttaagaacaacaaccccttca 1546
Db 1382 AGCACACTTTCAACCTTAAAAGCCACAGAGCCAGGCCAGGATAAAGATAACAAACCCCTTCA 1441
QY 1547 agttcttcgtccaccttttcagcagtcctccgagggcgtgtggctta 1592
Db 1442 AGTTCCTGTTCCACCTCAGGAGTCTCAGAAGAGAGCTGTGCTTAA 1487

RESULT 14
US-60-312-544-1040
; Sequence 1040, Application US/60312544
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)A
; CURRENT APPLICATION NUMBER: US/60/312,544
; CURRENT FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 10730
; SEQ ID NO 1040
; LENGTH: 1693
; TYPE: DNA

; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27)..(1484)
; OTHER INFORMATION: Clone ID: 700676766_FLI
US-60-312-544-1040

Query Match 27.6%; Score 510.8; DB 64; Length 1693;
Best Local Similarity 61.4%; Pred. No. 2e-78;
Matches 985; Conservative 0; Mismatches 487; Indels 132; Gaps 5;

QY 17 agctttcttttgccttttgccttttagttctgtggagctagcagcatctccttcaggcagc 76
Db 34 agcttgttcttccctttgttctccttcttttcag--tggctgcttgcctctgagagagc 90
QY 77 agccggaggagaaatgcgtgccagttccagcgcctcaatgcgcagagacacctgacaaccgca 136
Db 91 aggcacagcaaaatgagtgccagatccaaaagctgaatgccctcaaacccggataaccgta 150
QY 137 ttgaatcggaggcggttacattgagaccttggaaccccccaacaaaccaggaggttcgaatgcg 196
Db 151 tagagtcggaaggtggttcattgagacctggaacccctaacaacaagccattccagtggtg 210
QY 197 ccggcgtgcgcctctctcgtcttagtcctccgcgcgaacgcccttcgttaggcctttctact 256
Db 211 ccggtgttgccctctctcgtgcaccttaacgcgcaatgcccttcgtagaccttcctaca 270
QY 257 ccaatgctccccaggagatcttcattccagcaaggaaggggatactttgggttgatatcc 316
Db 271 ccaacggtccccaggaaatctacatacaacaaggtaatggtatttttggcatgatattcc 330
QY 317 ctggtgtccttagcacatatgaagagcctgcacaacaaggacgccgatatcagtcccaaa 376
Db 331 cggtgtgtccttagcacttatcaagagccgcgaagaa----- 365
QY 377 gaccaccaagacgttttgcaagaagaagaccaaaagccaaacagcaacaagatagtcaccaga 436
Db 366 -----tctcagcaacgaggacgaagccagagggcccaagaccgtcaccaaa 411
QY 437 aggtgcaccgttttcaatgaggggtgatctcattgcagttccccaccggtgtgttcttctggc 496
Db 412 aggtacatcgcttcagagaggggtgatttgatgcagctgcctactggtgtgcatggtgga 471
QY 497 tgtacaacgacacgacactgatgttgtgtgttttcttcttactgacaccaacaacaacg 556
Db 472 tgtacaacaatgaagacactcctgtgtgttcggtttctattattgacaccaacagcttgg 531
QY 557 acaaccagcttgatcagttccccaggagattcaatttggctgggaaccacgagcaagagt 616
Db 532 agaaccagctcgaccagatgcctaggagattctatcttctgtgggaaccaagagcaagagt 591
QY 617 tcttaaggtaccagcaacaagcagacaaagcagacgaagaagccttaccatatagcccat 676
Db 592 ttctaaaaatatcagcagca----- 610
QY 677 acagcccgcatagtcggccttagacgagaagagcgtgaatttcgccctcgaggacagcaca 736
Db 611 -----gcagcaagaggaggttccc 627
QY 737 gccgcagagaacgagcaggacaagaagaagaagacgaaggttggaacaatcttccagcggct 796
Db 628 aaagccagaagaagaaagcaacaagaagaagaaaaacgaaggaagcaacatatattgagtggt 687
QY 797 tcacgccggagttccttggaacaagccttcagggttgacgacagacagattgtgcaaaatc 856
Db 688 tcgccccctgaattcttgaaagaagcgttc---ggcgtgaacatgcagatagtgagaaacc 744
QY 857 tgtggggcgagaacgagagtgaaagaaggggagccatttgtacggtgaggggagcctca 916
Db 745 tacaaggtgagaacgaagagggagtagtgagccatttgtgacagtgaaaggaggtctaa 804
QY 917 gaatcttgagcccagatggaaacgagaggtgtccgacgaagaaggaatacagatgaagatc 976

| | | | |
|----|------|---|------|
| QY | 737 | gccgcagagaacgagcaggacagaagaagaagacgaaggtggaaacatcttcacgcgct | 796 |
| Db | 602 | AAAGCCAGAAGGAAAGCAACAAGAAGAAGAAAACGAAGGRAGCAACATATTGAGTGGCT | 661 |
| QY | 797 | tcacgcggagttccctggacaagccttccaggttgacgacagacagatttgtcaaaaac | 856 |
| Db | 662 | TCGCCCTGAATCTTTGAAAGAACGGTTC--GGCGTGAAATGCAGATAGTGAGAAACC | 718 |
| QY | 857 | tgtggggcgagaacgagagtgaagaagagggagccatttgtacggtgaggggagggcctca | 916 |
| Db | 719 | TACAAGGTGAGAACGAAGAGGAGGATAGTGGAGCCATTGTGACAGTGAAAGGAGGTCTAA | 778 |
| QY | 917 | gaatcttgagcccagatggaacgagaggtgccgacgaagaagaggaatacgcataagatc | 976 |
| Db | 779 | GAGTCACAGCTCCAGCCATGAGGAAGCCACAGCAAGAAAGATGATGATGATGAGGAAG | 838 |
| QY | 977 | aatatgaataaccatgaacaggatggaaggcgtggcagggg--aagcagagggcggggga | 1033 |
| Db | 839 | AGCAGCCACAGTGGTGGAGACAGACAAAGGTTGCCAACGCCAAAGCAAAGGAGGACAGAA | 898 |
| QY | 1034 | atggtattgaagagacgatctgcaccgcgatgtgttaaaaaaacattggtggaaacagat | 1093 |
| Db | 899 | ATGGCATTTGATGAGACCAATTTGCACAATGAGACTTCGCCAAAAACATTGGTCAGAATTCAT | 958 |
| QY | 1094 | cacctcacatctacgatcctcagcgcgtggttcactcaaaactgccacgatctcaaccttc | 1153 |
| Db | 959 | CACCTGACATCTACAACCCCTCAAGCTGGTAGCATCACAAACCCGCCACGCTTGAATTCC | 1018 |
| QY | 1154 | taatccttaggtggcttggacttagtgctgaatatggaaatctctacaggaatgcattgt | 1213 |
| Db | 1019 | CAGCCCTCTGGCTTCTCAAACTCAGTGCCCCAGTATGGATCACTCCGCAAGAATGCTATGT | 1078 |
| QY | 1214 | ttgtccctcaactacacaccaacgcacacagcatcataatatgcattgaggggagcggtc | 1273 |
| Db | 1079 | TCGTGCCACACTACACCCCTGAACGCGAACAGCAATAATATACGCATTGAATGGCGGGCAT | 1138 |
| QY | 1274 | acgtgcaagtggtggacagcaacggcaacagagtgtacgacgagagcttcaagagggtc | 1333 |
| Db | 1139 | TGGTACAAGTGGTGAAATGCAATGGTGAGAGAGTGTGTGATGGAGAGCTGCAAGAGGGAG | 1198 |
| QY | 1334 | acgttcttgtgtgocacagaacttcgccgtggtgggaagtcgccagagcgagaacttcg | 1393 |
| Db | 1199 | GGTGCTGATCGTTCCACAAAACTTTGCCGTGGCTGTCAAAAATCCCAGAGCGGATAACTTG | 1258 |
| QY | 1394 | aatacgtggcattcaagacagattcaaggccccagcatagccaactttgccggtgaaaact | 1453 |
| Db | 1259 | AGTATGTGTCAATTAAGACCAATGATAGACCCCTCGATCGGAACCTTGCAGGGGGCAACT | 1318 |
| QY | 1454 | ccttcatagataaacctgcggaggaggtggttgcaaatcattatggcctcccaaaggagc | 1513 |
| Db | 1319 | CATTGTTGAACGCATTGCCAGAGGAAGTGATTCAGCACACTTTTAACCTAAAGAGCCAGC | 1378 |
| QY | 1514 | aggcaaggcagcttaagaacaacaaccccttcaagttcttcgttccacccttttcagcagt | 1573 |
| Db | 1379 | AGCCAGGCAGGTGAAGAACAACAACCCCTTTCAGCTTCTTGTTCACCTCAGGAGTCTC | 1438 |
| QY | 1574 | ctccgagggctgtggtcta | 1592 |
| Db | 1439 | AGAGGAGAGCTGTGGCTTA | 1457 |

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 13:03:21 ; Search time 8470.79 Seconds
(without alignments)
3608.787 Million cell updates/sec

Title: US-09-715-036-4
Perfect score: 1853
Sequence: 1 atggctaagcttcttgagct.....aaaaaaaaaaaaaaaaaaaa 1853

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|------------|--------------------|
| 1 | 1853 | 100.0 | 1853 | 6 | AX148741 | AX148741 Sequence |
| 2 | 1853 | 100.0 | 1853 | 8 | AF086821 | AF086821 Arachis h |
| 3 | 1403.4 | 75.7 | 1590 | 8 | AF125192 | AF125192 Arachis h |
| 4 | 1391.2 | 75.1 | 1524 | 8 | AF093541 | AF093541 Arachis h |
| 5 | 1377.4 | 74.3 | 1524 | 6 | AX155334 | AX155334 Sequence |
| 6 | 516 | 27.8 | 1743 | 6 | E32414 | E32414 Transgenic |
| 7 | 516 | 27.8 | 1743 | 8 | SOYGLYBSU | M36686 Soybean gly |
| 8 | 512.4 | 27.7 | 1708 | 8 | SOYGA2B1A | D00216 Glycine max |
| 9 | 510.6 | 27.6 | 1712 | 6 | E02462 | E02462 cDNA encodi |
| 10 | 510.6 | 27.6 | 1712 | 8 | GMGLY1 | X02806 Glycine max |
| 11 | 508 | 27.4 | 1746 | 6 | E02463 | E02463 cDNA encodi |
| 12 | 508 | 27.4 | 1746 | 8 | GMGLY1A | X02985 Soybean mRN |
| 13 | 493.4 | 26.6 | 1729 | 8 | PSA132614 | AJ132614 Pisum sat |
| 14 | 449.6 | 24.3 | 1664 | 8 | VSLEGUMA | Z32835 V.sativa mR |
| 15 | 439.8 | 23.7 | 1666 | 8 | VFLEA2 | X55014 Vicia faba |
| 16 | 439.4 | 23.7 | 1650 | 8 | VNPROLEGA | Z46803 V.narbonens |
| 17 | 389.4 | 21.0 | 1585 | 8 | VFLEA1 | X55013 Vicia faba |
| 18 | 257.6 | 13.9 | 4156 | 8 | PSLEGA2 | X17193 Pisum sativ |
| 19 | 255 | 13.8 | 3347 | 8 | PSLEGAG | X02982 Pea lega ge |
| 20 | 242 | 13.1 | 787 | 8 | PEALGN | J01255 Pea (P.sati |
| 21 | 235.8 | 12.7 | 1787 | 8 | PABTPRU2 | X78120 P.amygdalus |
| 22 | 229.4 | 12.4 | 1784 | 8 | GSYG5 | X79467 G.soja (SH1 |
| 23 | 226.2 | 12.2 | 1755 | 8 | AB049440 | AB049440 Glycine m |
| 24 | 223.4 | 12.1 | 1763 | 8 | COTSPD | M16905 G.hirsutum |
| 25 | 223 | 12.0 | 1782 | 8 | AB000169 | AB000169 Glycine m |
| 26 | 221.4 | 11.9 | 1789 | 8 | AB000168 | AB000168 Glycine m |
| 27 | 208 | 11.2 | 3573 | 8 | GMGY3 | X15123 Soybean Gy3 |
| 28 | 205.2 | 11.1 | 1905 | 8 | PABTPRU1 | X78119 P.amygdalus |
| 29 | 204 | 11.0 | 1607 | 8 | SOYGLY | K02646 Soybean gly |
| 30 | 204 | 11.0 | 4617 | 8 | GMGY2 | X15122 Soybean Gy2 |
| 31 | 202.2 | 10.9 | 3176 | 8 | GMGLYCAB | Y00398 Soybean DNA |
| 32 | 201.8 | 10.9 | 1786 | 6 | E00868 | E00868 cDNA encodi |
| 33 | 201.8 | 10.9 | 1786 | 6 | E02460 | E02460 cDNA encodi |
| 34 | 201.8 | 10.9 | 1786 | 8 | SOYGLYAB | M10962 Soybean (G. |
| 35 | 201.6 | 10.9 | 3527 | 8 | GMGY1 | X15121 Soybean Gy1 |
| 36 | 195.2 | 10.5 | 1692 | 8 | QRRNALP | X99539 Q.robur mRN |
| 37 | 187.8 | 10.1 | 2282 | 8 | CAY15527 | Y15527 Cicer ariet |
| 38 | 187.6 | 10.1 | 1674 | 8 | AF319777 | AF319777 Glycine m |
| 39 | 173.4 | 9.4 | 1605 | 8 | VSLEGUMB | Z32796 V.sativa mR |
| 40 | 169.8 | 9.2 | 1950 | 8 | GSGLYCGY4 | X86970 G.soja mRNa |
| 41 | 168.2 | 9.1 | 1899 | 6 | E02461 | E02461 cDNA encodi |
| 42 | 168.2 | 9.1 | 1899 | 8 | GMGLYR1 | X02626 Soybean mRN |
| 43 | 162.6 | 8.8 | 1939 | 8 | SSPPLEGUM | Y09116 S.sagittifo |
| 44 | 160.8 | 8.7 | 2826 | 8 | HNNHAG3D1S | M28832 Sunflower 1 |
| 45 | 159.4 | 8.6 | 1679 | 8 | MSLEG11 | X82463 M.salicifol |

ALIGNMENTS

RESULT 1

AX148741

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

peanut.

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;

Aeschynomeneae; Arachis.

1 (bases 1 to 1853)

Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.

Down-regulation and silencing of allergen genes in transgenic

peanut seeds

Patent: WO 0136621-A 4 25-MAY-2001;

Alabama A & M University (US)

Location/Qualifiers

PAT 08-JUN-2001

AX148741 1853 bp DNA
Sequence 4 from Patent WO0136621.

AX148741
AX148741.1 GI:14347295

| | | | |
|------------|---|--|------|
| QY | 792 | cggcttcacgcgcggagttccttggaacaaagccttccagggttgacagacagacagattgtgca | 851 |
| Db | 723 | CGGCTTCACGCCGGAGTTCCTGGAACAGCCTTCAGGTTGACGACAGACAGATAGTGCA | 782 |
| QY | 852 | aaatctgtgggcgcgagaaacagagtgaagaagaggagccattgtgacggtgaggggagg | 911 |
| Db | 783 | AAACCTAAGAGCGGAGACTGAGAGTGAAGAAGAGGGAGCCATTGTGACAGTGAGGGGAGG | 842 |
| QY | 912 | cctcagaatcttgagcccgatggaacgagaggtgccgacgagagtggaagaagaggaaatacagatga | 971 |
| Db | 843 | CCTCAGAACTTGAGCCCGATAGAAAGAGACGTCGCCGACGAAGAAGAGGAATACGATGA | 902 |
| QY | 972 | agatcaaatatgaataccatgaacaggatggaaggcgtggcagggggaagcagagggcg9ggg | 1031 |
| Db | 903 | AGATGAATATGAATACGATGAAGAGGATAGAAGGCGTGGCAGGGGAAGCAGAGGCAGGGG | 962 |
| QY | 1032 | gaatggtatgaagagacgatctgcaccgcgatgtgttaaanaagaacattggtggaacag | 1091 |
| Db | 963 | GAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAAAGAACATTGGTAGAAACAG | 1022 |
| QY | 1092 | atccccacatctacgatcctcagcgctggtttcactcaaaaactgcc-acgatctcaacc | 1150 |
| Db | 1023 | ATCCCCTGACATCTACAACCCCTCA-AGCTGTTCACTCAAAACTGCCAACGATCTCAACC | 1081 |
| QY | 1151 | ttctaataccttagtggtgacctgactagtgctgaatatgaaatctctacaggaatgcac | 1210 |
| Db | 1082 | TTCTAATACTTAGTGGCTTGGACCTAGTGCTGAATATGAAATCTCTACAGGAATGCAT | 1141 |
| QY | 1211 | tgttgtccctcactacaacaccaacgcacacagcatcatatatatgcattgaggggacggg | 1270 |
| Db | 1142 | TGTTTGTGCTCACTACAACACCAACGGCACACAGCATCATATATCGATTGAGGGGACCGG | 1201 |
| QY | 1271 | ctcagctgcaagtgttgacagcaacggcaacagagtgctacgacgaggagcttcaagagg | 1330 |
| Db | 1202 | CTCACGTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGG | 1261 |
| QY | 1331 | gtcacgttcttgtggtgccacagaacttcgcgctggtggggaagtccccagagcgagaact | 1390 |
| Db | 1262 | GTCACGTGCTTGTGTTGCCACAGAACTTCGCCGTCGCTGGAAAGTCCCAGAGCGAGA | 1321 |
| QY | 1391 | tcgaatacgtggcattcaagacagattcaaggccccagcatagccaaactttgcggtgaaa | 1450 |
| Db | 1322 | TCGAATACGTGGCATTCGAAGACAGACTCAAGGCCCCAGCATAGCCAACTCGCCGGTG | 1381 |
| QY | 1451 | actccttcatatagataaacctgcgcgaggaggtggttgcaaaattcatatggcctcccaa | 1510 |
| Db | 1382 | ACTCCGTCATAGATAACCTGCGGGAGGAGGTGTTGCAAAATTCATATGGCCCTCCAA | 1441 |
| QY | 1511 | agcaggcaaggcagcttaagaacaacaacccccctcaagttcttcgttccacctttcagc | 1570 |
| Db | 1442 | AGCAGGCAAGGCAGCTTAAGAACAACAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG | 1501 |
| QY | 1571 | agtcctccgaggcgtgtggttaa | 1593 |
| Db | 1502 | AGTCTCCGAGGGCTGTGGCTTAA | 1524 |
| RESULT | 6 | | |
| E32414 | | | |
| LOCUS | E32414 | 1743 bp | DNA |
| DEFINITION | Transgenic plant with the expression of soybean glycinin. | | |
| ACCESSION | E32414 | | |
| VERSION | E32414.1 | GI:13022283 | |
| KEYWORDS | JP 2000050871-A/1. | | |
| SOURCE | Glycine max. | | |
| ORGANISM | Glycine max | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| | Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; | | |
| | Glycine. | | |
| REFERENCE | 1 (bases 1 to 1743) | | |
| AUTHORS | Fumio,T.S.U.U. and Katsube. | | |

| | | | |
|---|--|--|-------|
| TITLE | Transgenic plant with the expression of soybean glycinin | | |
| JOURNAL | Patent: JP 2000050871-A 1 22-FEB-2000; | | |
| | NATL INST OF AGROBIOLOGICAL RESOURCES | | |
| COMMENT | OS Glycine max (soybean) | | |
| | PN JP 2000050871-A/1 | | |
| | PD 22-FEB-2000 | | |
| | PF 07-AUG-1998 JP 1998223897 | | |
| | PR FUMIO TAKAIWA,SHIGERU UTSUMI,TOMOYUKI KATSUBE PC | | |
| | C12N15/09,A01H5/00,C12N5/10,C12P21/02//C07K14/415,(C12N15/09, PC | | |
| | C12R1:91), | | |
| | PC (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91), | | |
| | PC (C12N5/00,C12R1:91) | | |
| | CC | | |
| | FH Key | | |
| FEATURES | Key | Location/Qualifiers | |
| | CDS | (52)..(1536). | |
| | source | 1..1743 | |
| | | /organism="Glycine max" | |
| | | /db_xref="taxon:3847" | |
| BASE COUNT | 552 a | 419 c | 398 g |
| ORIGIN | 374 t | | |
| Query Match 27.8%; Score 516; DB 6; Length 1743; | | | |
| Best Local Similarity 62.3%; Pred. No. 5.1e-114; | | | |
| Matches 1002; Conservative 0; Mismatches 450; Indels 156; Gaps 6; | | | |
| QY | 17 | agctttcttttggcttttgccttcttagttctgtgggagctagcagcatctcctt-----ca | 70 |
| Db | 59 | AGCTAGTTTTTCCCTTTGTTTTCTGCTTTTCAGTGGCTGCTGCTTCGCTTTCAGTTCCA | 118 |
| QY | 71 | ggcagcagccggaggaagaatgcgtgcaggttccagcgcctcaatgcgcagagacctgaca | 130 |
| Db | 119 | GAGAGCAGCCTCAGCAAAACGAGTGCCAGATCCAAAACCTCAATGCCCTCAAACCGGATA | 178 |
| QY | 131 | accgcattgaatcggagggcggttacattgagacttggaaaccccaacacacaggagttcg | 190 |
| Db | 179 | ACCGTATAGAGTCAGAAGGAGGGCTCATTTGAGACATGGAACCCCTAACAAACAGCCATTCC | 238 |
| QY | 191 | aatgcgcgcgctgcctctctcgttagtacctcgcgcgaacgccttcgttaggcctt | 250 |
| Db | 239 | AGTGTGCCGGTGTGCCCCCTCTCTCGTGCACCCCTCAACCCGCAACGCCCTTCGTAGACCTT | 298 |
| QY | 251 | tctactccaatgctccccaggagatcttcatccagcaaggaaggggatactttgggttga | 310 |
| Db | 299 | CCTACACCAACGGTCCCCCAGGAATCTACATCCAAACAGGTAAAGGGTATTTTGGCATGA | 358 |
| QY | 311 | tattccctggttgctcctagcacatatgaagagcctgcacaacaagagcgccgatatcagt | 370 |
| Db | 359 | TATACCCGGGTTGTCTTAGCACATTGAAGAGCCTCAACAAAC----- | 400 |
| QY | 371 | cccaaaagaccaccaagacgttttgaagaagaagaccacaaagcaacaagatatagtc | 430 |
| Db | 401 | -----CTCAACAAAGAGGACAAAGCAGCAGACCACAAAGACCGTC | 439 |
| QY | 431 | accagaaggtgcaccgttttcaatgagggtgatctcattgcagttcccaccggtgttgcctt | 490 |
| Db | 440 | ACCAGAAGATCTATAACTTCAGAGAGGGTGATTGTATCGCAGTGCCCTACTGTTGTCAT | 499 |
| QY | 491 | tctggctgtacaacgaccacgacactgatgttgttgcgttttcttactgacaccaaca | 550 |
| Db | 500 | GGTGGATGTACAACAATGAAGACACTCCTGTTGTTGCCGTTCTATTATTGACACCAACA | 559 |
| QY | 551 | acaacgacaaccagcttgatcagttccccaggagagattcaatttggctgggaaccacgagc | 610 |
| Db | 560 | GCTTGAGAACCCAGCTCGACCCAGATGCCCTAGGAGATTCTATCTTCTGCTGGGAACCAAGAGC | 619 |
| QY | 611 | aagagttcttaagggtaccagcaacaagaagcagacaaaagcaggaagaagcgttaccatata | 670 |
| Db | 620 | AAGAGTTCTAAAAATATCAGCAAGA----- | 644 |
| QY | 671 | gcccatacagcccgcacatagtcgggcctagacgagagaagcgtgaatttcgccctcgaggac | 730 |


```
Db      645  -----GCAAGGAGGTC 655
QY      731  agcacagccgcagagaacgagcaggacaagaagaagacgaaggtggaacaacatctta 790
Db      656  ATCAAAGCCAGAAAGAAAGCATCAGCAAGAAGAAGAAACGAAGGAGGCAGCATATTGA 715
QY      791  gcggcttcacgcggagttccttggaacaagccttcagggttgacgacagacagattgtgc 850
Db      716  GTGGCTTCACCCCTGGAATTCTTGGAACATGCATTCAGCGT---GGACAAGCAGATAGCGA 772
QY      851  aaaaatctgtgggcgagagaacgagagtgaagaagagggagccattgtgacggtgaggggag 910
Db      773  AAAACCTACAAGGAGAGAACGAAGGGGAAGACAAAGGGGCCATTGTGACAGTGAAAGGAG 832
QY      911  gcctcagaatcttgagccca-----gatggaacgagaggtgcccgcgaag 955
Db      833  GTCTGAGCGTGATAAAACCACCCACGGACGAGCAGCAACAAAGACCCAGGAAGGAAG 892
QY      956  aagaggaatacgatgaagatcaata-----tgaataccatgaacaggatggaaggc 1006
Db      893  AAGAAGAAGAGGATGAGAAGCCACACAGTGCAGGGTAAAGACAAACACTGCCAACGCCCCC 952
QY      1007  gtggcaggggaagcadagggcgggggaatggtattgaagagacgatctgcaccgcgatgtg 1066
Db      953  GAGGAAGCCAAAGCAAAAGCAGAAGAAATGCGATTGACGAGACCATATGCACCATGAGAC 1012
QY      1067  ttaaaaaagaacattggtggaaacacagatccccctcacatctacgatcctcagcgctgttca 1126
Db      1013  TTCGCCACAACATTGGCCAGACTTTCATCACCTTGACATCTACAACCCCTCAAGCCGGTAGCG 1072
QY      1127  ctcaaaaactgccacgatctcaaccttttaaccttaggtggcttggaacttagtctgaat 1186
Db      1073  TCACAACCGCCACACGCGCTTGACTTCCCAGGCCCTCTCGTGGCTCAGACTCAGTGCTGAGT 1132
QY      1187  atggaaaatctctacaggaatgcattgttgttccctcaactacaacacccaacgcacacagca 1246
Db      1133  TTGGATCTCTCCGCAAGAATGCAATGTTTCGTGCCACACTACAACCTGAACCGGAACAGCA 1192
QY      1247  tcataatgcattgagggagcggtcacgtgcacgtggtggacagcaaacggcaacagag 1306
Db      1193  TAAATATACGCATTGAATGGACGGGCATTGTATACAAGTGGTGAATTGCAACGCTGAGAGAG 1252
QY      1307  tgtacgacgaggagcttcaagaggggtcacgttcttgtggtgccacagaaacttcgccgttg 1366
Db      1253  TGTTTGATGGAGAGCTGCAGAGGGGACGGGTGCTGTGATCGTGCCACAAAACCTTGTGGTGG 1312
QY      1367  ctgggaagtcccagagcgaagaacttcgaatacgtggttcattcaagacagattcaaggccca 1426
Db      1313  CTGCAAGATCACAGAGTGACAACTTCGAGTATGTGTGATTCAAAGACCAATGATACACCCA 1372
QY      1427  gcatagccaactttgccggtgaaaaactccttcatagataaacctgccgagggaggtggttg 1486
Db      1373  TGATCGGCACTCTTGCAAGGGGCCAAACTCATTTGTTGAACGCATTACCAGAGGGAAGTGATT 1432
QY      1487  caaatcokatatggcctcccaagggagcaggcaaggcaggttaagaacaacaaccccttca 1546
Db      1433  AGCACACTTTCAAACCTAAAGCCACGACGAGGCCAGGCAGACATAAAGAACACACCCCTTTCA 1492
QY      1547  agttcttggttccacaccttttcagcagtcctccgagggctgtggcttaaa 1594
Db      1493  AGTTCCTGGTTCACACTCAGGAGTCTCAGAAGAGAGAGCTGTGGCTTAGA 1540
```

```
RESULT 7
SOYGLYBSU      1743 bp      mRNA      PLN      27-APR-1993
LOCUS      Soybean glycinin A-la-B-x subunit mRNA, complete cds.
DEFINITION      M36686 D00566
ACCESSION      M36686.1 GI:169972
VERSION      glycinin; storage protein.
KEYWORDS      Soybean (var. Shirotsurunoko) cotyledon, cDNA to mRNA, clones
SOURCE      pGST[8,4-2-3-5,4-2-11-10].
```

```
ORGANISM      Glycine max
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
               Glycine.
REFERENCE      1 (bases 1 to 1743)
AUTHORS      Utsumi,S., Kohn,M., Mori,T. and Kito,M.
TITLE      An alternate cDNA encoding glycinin A-la-B-x subunit
JOURNAL      J. Agric. Food Chem. 35, 210-214 (1987)
FEATURES      Location/Qualifiers
               1..1743
               /organism="Glycine max"
               /db_xref="taxon:3847"
               /tissue_type="cotyledon"
               52..1539
               /note="glycinin A-la-B-x subunit"
               /codon_start=1
               /protein_id="AAA33966.1"
               /db_xref="GI:169973"
               /translation="MAKLVFSLCFLLFSGCCFAFSSREOPQONECQIQKLNALKPDNR
               IESEGLIETWNPNNKPFQCGVALSRCTLNRNALRRPSYTNQPGQEIYIQQKGIFGM
               IYPCPSTFEEPQPOQRGQSSRPDRHQKIYNFREGDLIAVPTGVAWWMYNNEDTPV
               VAVSIIDTNSLENQLDQMPRRFYLAGNQEFLKYQQEGGHQSQKHKHQEENEKG
               SILSGFTLEHAFSVDKQIAKNLQGENEGEDKGAIVTVKGLSVIKPPTDEQQORP
               QEEEEDEKPKQCKDKHKCQRPRGSQSKSRNRIDETICTMRLRHNIGQTSSPDIY
               NPQAGSVTATSLDFPALSWLRLSAEFGSLRKNAMFVPHYNLANSIIYALNGLRIQ
               VVNCNGRVFDGELQGRVLIVPQNFVVAARSQSDNFEYVSKTNDTPMIGTLAGANS
               LLNALPEVIQHTFNLKSQARQIKNNPFKFLVPPQESQKRAVA"
               polyA_signal      1597..1602
               polyA_signal      1711..1716
BASE COUNT      552 a      419 c      398 g      374 t
ORIGIN
```

```
Query Match      27.8%; Score 516; DB 8; Length 1743;
Best Local Similarity 62.3%; Pred. No. 5.1e-114;
Matches 1002; Conservative 0; Mismatches 450; Indels 156; Gaps 6;

QY      17  agctttcttttgcctttgcttagttctgtggagctagcagcatctcctt-----ca 70
Db      59  AGCTAGTTTTCCTTTTCTGCTTTTTCAGTGGCTGCTGCTTCGCTTCAGTTCCA 118
QY      71  ggcagcagccggaggagaatgcgtgccagttccagcgcctcaatgcgcagagacctgaca 130
Db      119  GAGAGCAGCCTCAGCAAAACAGTGCAGATCCAAAACCTCAATGCCCTCAAAACCGGATA 178
QY      131  accgcattgaatcggagggcggttacattgagacttggaaccccccaacacaggagttcg 190
Db      179  ACCGTATAGAGTCAGAAGGAGGGCTCATTTGAGACATGGAACCTTAACAAACAGCCATTCC 238
QY      191  aatgcgcggcgctgcgccctctctgcgttagtcctccgcgcgaacgcccttcctagccctt 250
Db      239  AGTGTGCCGCTGTGCCCTCTCTCGCTGCACCCCTCAACCGCAACGCCCTTCGTAGACCTT 298
QY      251  tctaactcaaatgctccccaggagatcttcatccagcaaggagggtgatactttgggttga 310
Db      299  CCTACACCAACGGTCCCCAGGAATCTACATCCAACAAGSTAAGGGTATTTTGGCATGA 358
QY      311  tattccctggttgctcctagcacatatgaagagcctgcacacaaggacgcgcgatatcagt 370
Db      359  TATACCCGGGTTGTCTCCTAGCACATTTGAAGAGCCTCAACAAC----- 400
QY      371  cccaaagaccacaagacggttttgaagaagaagaccaaaagccaacagcaacaagatagtc 430
Db      401  -----CTCAACAAAGAGGACAAAGCAGACACCACACAGACCCGTC 439
QY      431  accagaaggtgcaccggtttcaatgaggggtgatctcattgcagttccaccgggtgtgctt 490
Db      440  ACCAGAAGATCTATAAATTCAGAGAGGGTGATTTGATCGCAGTGCCCTACTGGTGTGCAT 499
QY      491  tctggctgtacaacgaccacgacacactgatgtgtgtgctgtttcttactgacaccaaca 550
Db      500  GGTGGATGTACAACAATGAAGACACTCCTGTTGTTGCCGTTTCTATTATTGACACCAACA 559
```


Db 518 GACATTAGAAGCTCCAATAACCAAGCTTGATCAGATGCCTAGGAGATTCTATCTTGCTGGG 577

QY 601 aaccacgagcaagagttctttaaggttacacgacaacaaagcagacaaagcagacgagcaagaagc 660

Db 578 AACCACGAGCAAGAGGTTTCTACAATACCAGCATCA---ACAAGGAGGAAAGCACAACA 633

QY 661 ttaccatatagcccatcacagcccgcatagtcggcctagacgagaagagcgtgaatttcgc 720

Db 634 AGAAAATGAAGGCAACAACATTTTCAGTGGCTTCAAGAGGGATTCTTGAAGATGCTTT 693

QY 721 cctcgaggacagcagccgcgagagaaacgagcaggacaaagaagaagacgaagtgga 780

Db 694 CAACGTGAACAGGCATATAGTAGACAGACTTCAAGGCAGGAATGAAGACGAAGAAGGG 753

QY 781 aacatcttcagcggcttcacgcccggagttcctggaacaagccttc-----aggttg 832

Db 754 AGCCATTGTCAAAGTGAAGGTTGGACTCAGCATCATAAAGCCACCCGAGAAAGCAGCGCG 813

QY 833 acgacagacagattgtgcaaaatctgtggggcgagaacgagagtggaagaagaggagcca 892

Db 814 CCACCAGAGAGGCAGACAGACAAGAGGAAGATGAAGATGAAGAGAGGCAGCCCGCGCACCA 873

QY 893 ttgtgacggtgaggggagccctcagaatcttgagcccgatggaacgagaggtgcgcagc 952

Db 874 GAGAGGCAGCAGACAAGAGGAAGAGGATGAAGATGAAGAGAGGCAGCCCGCTCATCA 933

QY 953 aagaagagga--atacgatgaagatcaatatgaataccatgaacaggatggaaggcgtg 1010

Db 934 AAGGAGAAGAGGAGAGGAAGAAAGACAAGAAAGAGCGCGCGGAGCCCAAAAAGG 993

QY 1011 caggggaaagcagagggggggaatggtattgaagagacgatcgcaccgcatgtgttaa 1070

Db 994 CAAAAGCAGAAGGCAGAGAGACAATGGGCTTGAGGAACAGTTTGCACTGCTAAACTCG 1053

QY 1071 aaagaacattggtggaacagatccccctcacatctacgatcctcagcgtggttcaactca 1130

Db 1054 ATTGAACATTGGCCCGCTTTCATCACCCAGACATCTACAACCCCTGAAGCTGGTAGAATCAA 1113

QY 1131 aaactgccacgatctcaaccttctaactccttaggtggttgacctagtcgtgaatatgg 1190

Db 1114 AACTGTTACCAAGCCTGGACCTCCCAGTTCTCAGGTGGCTCAAACCTGAAGTGTGAGCATGG 1173

QY 1191 aaatctctacaggaatgcattgtttgtccctcactacaacacacgaacgcacacagcatcat 1250

Db 1174 ATCTCTCCACAAAATGCTATGTTTGTGCTCACTACAACCTGAATGCAAAACAGTATAAT 1233

QY 1251 atatgcattgagggagcgggctcacggtgaaagtggtgacagcaacggcaacagagtgta 1310

Db 1234 ATACGCATTGAAGGCGAGCTGCAAGGCTACAAGTAGTGAACCTGCAATGGCAACACCCGTGT 1293

QY 1311 cgacgaggagcttcaagaggggtcacggttctgtggtgccacaagaacttcgcggtgctgg 1370

Db 1294 TGATGGAGAGCTAGAACCGGACGTCGATTCATATGACAGTGGCCACAAAACATATGCTGTGGCTGC 1353

QY 1371 gaagtccagagcgaagaacttcgaatacgtggcattcaagacagattcaaggccccagcat 1430

Db 1354 AAAGTCACTAAGCGACAGGTTCTCATATGTAGCATTCAAGACCAATGATAGAGCTGGTAT 1413

QY 1431 agccaaactttgccggtgaaaaactccttcatagataaacctgcccggaggaggtggttgcaaa 1490

Db 1414 TGCAAGACTTGCAGGGACATCATCAGTTATAAATAATCTGCCGTGGATGTTGTTGCAGC 1473

QY 1491 ttcatatgcccctcccaaggagcagggcagcttaagaacaacaaccccccttcaagtt 1550

Db 1474 TACATTCAACCTGCAGAGGAATGAGGCAAGCGAGCTCAAGTCCAACAATCCCTTCAAAAT 1533

QY 1551 cttcgttccaccttttcagcagtcctccgagggcgtgctggcttaaaaaacgaccagtatcttt 1610

Db 1534 TCTAGTTCAGCTCGTGAGTCTGAGAACACAGAGCTTCGGCTTAGATTGGCACCAACATCAA 1593

QY 1611 tgcaagcgtgttatccactaacataaactttttgcccacaaatgaataataataataaga 1670

Db 1594 TGAAGTAAATGAATAGAAAACCTAAGGCTTAGATGCCTTTTGTACTTGTGTAAATAACT 1653

QY 1671 agaataatgtagttttaatttttagtatgaataagaata 1709

Db 1654 CGAGTCATGTACCTTTTAGCGGAACACAGATAAATAAAA 1692

RESULT 14

VSLEGUMA

LOCUS VSLEGUMA 1664 bp mRNA PLN 03-MAY-1994

DEFINITION V.sativa mRNA for legumin A.

ACCESSION Z32835

VERSION Z32835.1 GI:483448

KEYWORDS l1s globulin; legumin A.

SOURCE .spring vetch.

ORGANISM Vicia sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.

REFERENCE 1 (bases 1 to 1664)

AUTHORS Nong,V., Becker,C. and Muentz,K.

TITLE Cloning and heterologous expression of cDNAs encoding legumins of vetch (Vicia sativa) seeds

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1664)

AUTHORS Nong,V.

TITLE Direct Submission

JOURNAL Submitted (28-APR-1994) Nong V., Institut fuer Pflanzenegenetik und Kulturpflanzenforschung, Molecular Cell Biology, Corrensstrasse 3, Gatersleben, Sachsen-Anhalt, Germany, D-06466

FEATURES

Location/Qualifiers

1..1664

/organism="Vicia sativa"

/db_xref="taxon:3908"

/clone="pVSLeA42 (1-1612), pVSLeA16 (461-1628), pVSLeA5 (961-1664)"

/tissue_type="Cotyledon"

/clone_lib="PCR-amplified cDNA"

/dev_stage="Seed, storage deposition stage"

1..1664

24..86

24..1520

/function="storage protein"

/citation=[1]

/codon_start=1

/evidence=experimental

/product="legumin A"

/protein_id="CAA83677.1"

/db_xref="GI:483449"

/db_xref="SPTREMBL:Q41702"

/translation="MAKLLALSFCFLLFSSCFALREQSQNECQLERINALEPDNR IESEGLIETWPNNRQRCARVALSRATLQRNALRPYYSNAPQEIYIQGNGYFGM VFGPCPETHPEEQSEGEGRYRDSHQVNRREGDIIAVPTGIAFWMYNDQDTPVI AISLDTGSSNNQLDQMPRRFYLAGNQEQEFLRYHQQGGKQEQDNDGNNTFSGFKRD FLEDAFNVRHIVDRLQRNEDEKGAIVKVGGLSIAPPERQARHERGSRQEEDED EKEERQPSHHKSRDEDEDDKEKRHSQKGQSRROGDNGLLEETVCTAKLRANIGSSPSP DIYNPOAGRIKTVTSLDLPVLRWLKLSAEHGS�HKNAMEVPHYNLNANSYIALKGRA RLQVVCNCGNTVFDGELEAGRALTVPQNYAVAAKSLSERFTYVAFKTDADRASLARLAG TSSVIDDLPLDVAATFNMQRNEARQLKSNNPFKFLVPPRQSEMRASA"

87..1517

/function="storage globulin"

/evidence=experimental

/product="legumin A; l1s globulin"

polyA_signal 1635..1645

BASE COUNT 538 a 358 c 383 g 385 t

ORIGIN

Query Match 24.3%; Score 449.6; DB 8; Length 1664;

Best Local Similarity 58.6%; Pred. No. 5.7e-98;

Matches 934; Conservative 0; Mismatches 564; Indels 96; Gaps 5;

QY 1 atggcctaagctcttgagcctttcttttgccttttgccttttgccttttgcctttgccttagcttgcgagctagcagc 60

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 13:04:00 ; Search time 8470.79 Seconds
(without alignments)
3957.396 Million cell updates/sec

Title: US-09-715-036-5
Perfect score: 2032
Sequence: 1 aataatcatatatattcatc.....cgtttgtcggtgtttctcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues 2944280

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htgo_hum:*
- 31: em_htgo_inv:*
- 32: em_htgo_rod:*
- 33: em_htg_hum:*
- 34: em_htg_inv:*
- 35: em_htg_rod:*
- 36: em_htg_Other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|--------|-------------|--------|----|-----------|--------------------|
| 1 | 2032 | 100.0 | 2032 | 6 | AX148742 | AX148742 Sequence |
| 2 | 2032 | 100.0 | 2032 | 8 | ARQARAH1 | L34402 Arachis hyp |
| 3 | 2019.2 | 99.4 | 2032 | 6 | AX155331 | AX155331 Sequence |
| 4 | 1805.2 | 88.8 | 1978 | 6 | AX155332 | AX155332 Sequence |
| 5 | 1752.4 | 86.2 | 1949 | 8 | ARQARAH | L38853 Arachis hyp |
| 6 | 465.8 | 22.9 | 1433 | 8 | PSVIC | Y00722 Pisum sativ |
| 7 | 461 | 22.7 | 1561 | 8 | VFVIC | Y00462 Vicia faba |
| 8 | 449 | 22.1 | 1596 | 8 | PSVICK | X67429 Pisum sativ |
| 9 | 435.4 | 21.4 | 1548 | 8 | VNVICLN | Z71987 V.narbonens |
| 10 | 408.4 | 20.1 | 1251 | 8 | AB008679 | AB008679 Glycine m |
| 11 | 404 | 19.9 | 1632 | 8 | AB008678 | AB008678 Glycine m |
| 12 | 404 | 19.9 | 1955 | 8 | GMASBC | X17698 Soybean mRN |
| 13 | 402 | 19.8 | 1680 | 8 | AB008680 | AB008680 Glycine m |
| 14 | 378.8 | 18.6 | 2069 | 8 | PSA276875 | AJ276875 Pisum sat |
| 15 | 347.8 | 17.1 | 1858 | 8 | VNCONVN | Z71986 V.narbonens |
| 16 | 339.4 | 16.7 | 1494 | 8 | GMU14006 | U14006 Glycine max |
| 17 | 336.2 | 16.5 | 1086 | 8 | PEAVIC7 | J01260 pea vicilin |
| 18 | 304.6 | 15.0 | 1534 | 8 | CGCANAV | X06733 Canavalia g |
| 19 | 303.2 | 14.9 | 907 | 8 | PEAVIC2 | J01258 pea vicilin |
| 20 | 303 | 14.9 | 1403 | 8 | CECANAV | X59467 C.ensiformi |
| 21 | 250.4 | 12.3 | 1035 | 8 | SOY7SAA | J01290 soybean 7s |
| 22 | 239.6 | 11.8 | 1047 | 8 | SOY7SAPA | J01294 soybean 7s |
| 23 | 204.4 | 10.1 | 794 | 8 | SOYBCONGL | M26128 Soybean bet |
| 24 | 178 | 8.8 | 530 | 8 | SOY7SAB | J01291 soybean 7s |
| 25 | 178 | 8.8 | 936 | 8 | SOY7SAC | J01292 soybean 7s |
| 26 | 175.4 | 8.6 | 1404 | 8 | U01122 | U01122 Phaseolus l |
| 27 | 174.2 | 8.6 | 1535 | 8 | PVPHASAR | X02980 Phaseolus v |
| 28 | 172.2 | 8.5 | 1478 | 8 | U01121 | U01121 Phaseolus l |
| 29 | 171.4 | 8.4 | 1741 | 8 | LCVICNE | Z48440 L.culinaris |
| 30 | 167 | 8.2 | 1341 | 8 | AF348366 | AF348366 Cajanus c |
| 31 | 160.6 | 7.9 | 591 | 8 | PSCONVR1 | X01379 Pea mRNA fr |
| 32 | 160.2 | 7.9 | 2140 | 6 | AX031079 | AX031079 Sequence |
| 33 | 160.2 | 7.9 | 2141 | 8 | AF161885 | AF161885 Macadamia |
| 34 | 160 | 7.9 | 1742 | 8 | LCVICNA | Z48436 L.culinaris |
| 35 | 160 | 7.9 | 1743 | 8 | LEVICN | Z48438 L.ervoides |
| 36 | 157.4 | 7.7 | 2171 | 6 | AX031077 | AX031077 Sequence |
| 37 | 157.4 | 7.7 | 2182 | 8 | AF161884 | AF161884 Macadamia |
| 38 | 157 | 7.7 | 4239 | 8 | PSVICIL | X14076 Pea vicilin |
| 39 | 155.6 | 7.7 | 1475 | 6 | A11822 | A11822 Artificial |
| 40 | 154 | 7.6 | 1454 | 8 | U01131 | U01131 Phaseolus v |
| 41 | 154 | 7.6 | 1475 | 8 | PVPHASBR | X03004 Phaseolus v |
| 42 | 154 | 7.6 | 5327 | 8 | VFVICG | Y00506 Vicia faba |
| 43 | 152.8 | 7.5 | 1430 | 8 | U01132 | U01132 Phaseolus v |
| 44 | 152.6 | 7.5 | 2171 | 6 | AX031075 | AX031075 Sequence |
| 45 | 152.6 | 7.5 | 2182 | 8 | AF161883 | AF161883 Macadamia |

ALIGNMENTS

| | | | | | |
|------------|-----------------------------------|-------------|-----|-----|-------------|
| RESULT | 1 | | | | |
| AX148742 | | | | | |
| LOCUS | AX148742 | 2032 bp | DNA | PAT | 08-JUN-2001 |
| DEFINITION | Sequence 5 from Patent WO0136621. | | | | |
| ACCESSION | AX148742 | | | | |
| VERSION | AX148742.1 | GI:14347296 | | | |
| KEYWORDS | | | | | |
| SOURCE | peanut. | | | | |
| ORGANISM | Arachis hypogaea | | | | |

REFERENCE 1 (bases 1 to 2032)
AUTHORS Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.
TITLE Down-regulation and silencing of allergen genes in transgenic peanut seeds
JOURNAL Patent: WO 0136621-A 5 25-MAY-2001;
FEATURES Alabama A & M University (US) Location/Qualifiers

Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 2032)

X

[illegible]

RESULT 2

ARQARAHI

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

2032 bp

Arachis hypogaea (clone P41b)

GI:602435

L34402

L34402.1

allergen.

Arachis hypogaea (strain Florunner).

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

PLN

Ara h I mRNA, complete cds.

24-MAY-1996

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

1 (bases 1 to 2032)

Burks,A.W., Cockrell,G., Stanley,J.S., Helm,R.M. and Bannon,G.A.

Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity

J. Clin. Invest. 96 (4), 1715-1721 (1995)

96013631

On Dec 16, 1994 this sequence version replaced gi:508640.

Location/Qualifiers

1..2032

/organism="Arachis hypogaea"

/strain="Florunner"

/db_xref="taxon:3818"

/clone="p41b"

/dev_stage="seed"

<1..49

50..1930

/gene="Ara h I"

50..1930

/gene="Ara h I"

/codon_start=1

/protein_id="AAB00861.1"

/db_xref="GI:602436"

/translation="MRGRVSPMLLLGILVLASVSATHAKSSPYQKKTENPCARCLQSCQEPDDLKQKACESRCTKLEYDPRCVYDPRGHTGTNORSPPGERTRGROPQDYDDRRQPRREEGRWGPAGPREREEDWRQPREDRWRPSSHQPRKIRPEGEQEWEWGT PGSHVREETSRNNPFYFPRSRRFSTRYGNQNGRIRVLRQFDQRSRQFNLRNHRIVQIE AKPTLVLPKHADADNLV IQQOATVTVANGNNRKSFNLDDEGHALRIPSGFISYILN RHDNQLRVAKISMPVNTPGQFEDFFPASSRDQSSVYLGFSRNTLEAAFNAAFNEIRR VLLENAGGEQEEERGQRWRSTRSENNEGVIIVKVSKEHVEELTKHAKSVSKKGSSEEG DITNPINLREGEPLDNNFGKLFVQKDPKKNPQLQDLDMMLTCVEIKEGALMLPHFNS KAMVIVVNVKGTGNELVAVRKEQQQRGRREEDEDEEEGSRREVRRTARLKEGD VFIMPAAHPVA INASSELHLLGFGINAEENHRIFLAGDKDNVIDQIEKAKDLAFPGS GEQVEKLIKQKESHFVSARPSQSQSPSPSEKESPEKEDQEEENQGGKPLLSILKA FN"

3'UTR

polyA_signal

polyA_site

BASE COUNT

ORIGIN

1931..2032

2005..2010

2032

628 a

473 c

530 g

401 t

Query Match

Best Local Similarity

Matches 2032; Conservative

100.0%;

100.0%;

0;

Score 2032;

DB 8;

Length 2032;

Indels 0;

Gaps 0;

QY

1

aataatcatatattcatcaatcatctatataagtagtagcaggagcaatgagagggag

60

Db

1

AATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGGAG

60

QY

61

ggtttctccactgatgctgttgctagggatccttgctcctggcttcagtttctgcaacgca

120

Db

61

GGTTTCTCCACTGATGCTGTGTGCTAGGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGCA

120

QY

121

tgccaagtcatcaccttaccagaagaacacagagaacccctgcgccagaggtgcctcca

180

Db

121

TGCCAAGTCATCACCTTACCAGAAGAAAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCA

180

QY

181

gagttgtcaacagggaacccgatgactgaagcaaaaggcatgctgcgtgcaccaa

240

181

Db

GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCA

240

241

QY

gctcgagtatgatcctcgttgtgtctatgatcctcgcaggacacactggcaccaccaacca

300

241

Db

GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCA

300

301

QY

acgttccctccaggggagcgacacagtggccgccaacccggagactacgatgatgaccg

360

301

Db

ACGTTCCCTCCAGGGGAGCGGACACAGTGGCCGCCAACCCGGAGACTACGATGATGACCG

360

361

QY

ccgtcaacccccgaagagaggaaggagccgatggggaccagctggaccgagggagcgtga

420

361

Db

CCGTCAACCCCGAAGAGAGAGGAAGGAGGCCGATGGGGACCAGCTGGACCGGAGCGCTGA

420

421

QY

aagagaagaagactggagacaaccaagaagaagattggaggcgaccaagtcatcagcagcc

480

421

Db

AAGAGAAGAAGACTGGAGACACCAACCAAGAGAAGATTGGAGCGGACCAAGTTCATCAGCAGCC

480

481

QY

acggaaataaagcccccgaaggaagagaagagaacaagatgggggaacaccaggttagcca

540

481

Db

ACGGAATAAAGGCCCGGAAGGAAGAGAAGAGAAACAAGAGTGGGGAACACCAGGTAGCCA

540

541

QY

tgtgaggggaagaaacatctcggaaacacccctttctactctccgtcaaggcggttagcac

600

541

Db

TGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGGCGGTTAGCAC

600

601

QY

ccgtacgggaaccaaaccggttaggattccgggtcctctgcagaggtttgaccaaagggtcaag

660

601

Db

CCGCTACGGGAACCAAAACCGGTAGGATCCGGGTCTCTGCAGAGGTTTGACCAAAAGGTCAAG

660

661

QY

gcagtttcagaatctccagaatcacccgtattgtgcagatcgaggccaaacctaactct

720

661

Db

GCAGTTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT

720

721

QY

tgttctcccaagcagcgtgatgctgataaacatccttcttatccagcaaggcgcaagccac

780

721

Db

TGTTCTTCCCAAGCACGCTGATGCTGATAAACATCCTTGTATTCCAGCAAGGGAAGCCAC

780

781

QY

cgtgaccgtagcaaatggcaataacagaaaagagctttaactcttgacgagggccatgcact

840

781

Db

CGTGACCGTAGCAAAATGGCAATAACAGAAAAGAGCTTTAATCTTGACGAGGGCCATGCCT

840

841

QY

cagaatcccatcccggtttcattcctacatcttgaaccgcatgacaaaccagaacctcag

900

841

Db

CAGAAATCCCATCCGGTTTCATTTCTTCTACATCTTGAAACCGCCATGACAACCAACCTCAG

900

901

QY

agtagctaaaaatctccatgccccgttaacacacacccggccagtttgaggattttctcccgcc

960

901

Db

ACTAGCTAAAATCTCCATGCCCGTTAACACACCCCGGCCAGTTTGAGGATTCTTCCCGGC

960

961

QY

gagcagccgagaccaatcatcctacttgagggttcaggggttcagcaggaatacgttgaggccgc

1020

961

Db

GAGCAGCCGAGACCAATCATCTCTACTTTCAGGGGCTTCAGCAGGAATACGTTGGAGGCCGC

1020

1021

QY

cttcaatgcggaattcaatgagatgagatcgagggtgctgttagaagagaatcagagggtga

1080

1021

Db

CTTCAATGCGGAATTCAATGAGATACGAGGGTGTGTGTAGAAGAGAATCGAGGAGGTGA

1080

1081

QY

gcaagaggagagaggcgagagcgcatggagtagtactcggagtagtgagaaacaatgaaggagt

1140

1081

Db

GCAAGAGGAGAGAGGCGCAGAGCGGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGT

1140

1141

QY

gatagtcaaatgtcaaaaggagcagcttgagaacttactaagcagcgtctaaatccgtctc

1200

1141

Db

GATAGTCAAAGTGTCAAAGGAGCACCGTTGAAGAACTTACTAAGCACCGCTAAATCCGTCTC

1200

1201

QY

aaagaaaggctccgaagaagaggagagatatcaaccaaccccaatcaacttgagagaagcgga

1260

1201

Db

AAAGAAAGGCTCCGAGAGAGAGGGAGATATCACCAACCCCAATCAACTTGAGAGAAGGCCGA

1260

1261

QY

gccccgatctttctaaacactttgggaagttatttgaggtgaagccagacaagaagaccc

1320

1261

Db

gccccgatctttctaaacactttgggaagttatttgaggtgaagccagacaagaagaccc

1320

✱

| | | | |
|------------|---|---|-------------|
| Db | 1261 | GCCCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCC | 1320 |
| Qy | 1321 | ccagcttcaggacacgtgacatgatgctcacctgtgtagagatcaaaagaaggcctttgat | 1380 |
| Db | 1321 | CCCCGATCTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAAGGAGCTTGAT | 1380 |
| Qy | 1381 | gctccacacttcaactcaaaagcccatggttatcgtctgctcaacaaaaggaactggaaa | 1440 |
| Db | 1381 | GCTCCACACTTCAACTCAAAAGGCCATGTTATCGTCGTCGTCAACAAAAGGAAGTGGAAA | 1440 |
| Qy | 1441 | ccttgaactcgtggctgtaagaaagagcaacaacagaggggacgcggggaagagagga | 1500 |
| Db | 1441 | CCTTGAACCTCGTGGCTGTAGAAAAGAGCAACAACAGAGGGACGGCGGGAAGAGGA | 1500 |
| Qy | 1501 | ggacgaagacgaagaagagggaagtaacagagaggtgctgtaggtacacagcgaggtt | 1560 |
| Db | 1501 | GGACGAAGACGAAGAAGAGGGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT | 1560 |
| Qy | 1561 | gaaggaagcgcatgtgttctcatcctgcccagcagctcatccagtagccatcaacgcttcctc | 1620 |
| Db | 1561 | GAAGGAAGCGCATGTGTTTCATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC | 1620 |
| Qy | 1621 | cgaactccatctgcttgcttgcttcggtatcaacgctgaaaaacaacacagaaatcttccttgc | 1680 |
| Db | 1621 | CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAACAACACAGAAATCTTCCCTTGC | 1680 |
| Qy | 1681 | aggtgataagacaatgtgatagacagatagagaagcaagcaggaagatttagcattccc | 1740 |
| Db | 1681 | AGGTGATAAGGACAAATGTATAGACAGATAGAGAAGCAAGCGAAGGATTTAGCAATCCC | 1740 |
| Qy | 1741 | tgggtcgggtgaacaagttgagaagctcatcaatactccgtcgtctcctgaaaagagtcctccttgtag | 1800 |
| Db | 1741 | TGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAAACCAGAAAGGAATCTCACTTTGTGAG | 1800 |
| Qy | 1801 | tgctcgtcctcaatctcaatctcaatctccgtcgtctcctgaaaagagtcctcctgagaa | 1860 |
| Db | 1801 | TGCTCGTCCCTCAATCTCAATCTCAATCTCCGTGCTCTCCTGAGAAAGAGTCTCCTTGAGAA | 1860 |
| Qy | 1861 | agaggatcaagaggaggaaaaccaaagggaagggtccactccttcaattttgaaggc | 1920 |
| Db | 1861 | AGAGGATCAAGAGGAGGAAACCAAGGAGGAAGGGTCCACTCTTCAATTTTGAAGSC | 1920 |
| Qy | 1921 | ttttaactgagaatggaggcaacttgttatgtatcgataataagatcacgcctttgtact | 1980 |
| Db | 1921 | TTTTAACTGAGAAATGGAGGCAACTTGTTATGTATCGATAATAGATCACGCTTTGTACT | 1980 |
| Qy | 1981 | ctactatccaaaacttatcaataataataaaaacggttgtgctgtgttctctcc | 2032 |
| Db | 1981 | CTACTATCCAAAACCTTATCAATAATAATAAAAACGTTTGTGCGTGTGTTCTCC | 2032 |
| RESULT | 3 | | |
| AX155331 | | | |
| LOCUS | AX155331 | 2032 bp | DNA |
| DEFINITION | Sequence 4 from Patent WO0140264. | PAT | 22-JUN-2001 |
| ACCESSION | AX155331 | | |
| VERSION | AX155331.1 | GI:14536766 | |
| KEYWORDS | | | |
| SOURCE | peanut. | | |
| ORGANISM | Arachis hypogaea | | |
| | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| | Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; | | |
| | Aeschynomeneae; Arachis. | | |
| REFERENCE | 1 (bases 1 to 2032) | | |
| AUTHORS | Bannon,G.A., Burks,W.A., Caplan,M.J., Sampson,H. and Sosin,H. | | |
| TITLE | Peptide antigens | | |
| JOURNAL | Patent: WO 0140264-A 4 07-JUN-2001; | | |
| | Panacea Pharmaceuticals, LLC (US) ; The University of Arkansas (US) | | |
| | ; MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK | | |
| | (US) | | |
| FEATURES | Location/Qualifiers | | |
| source | 1. .2032 | | |

| | | | | | |
|----|-----------------------|--|---------------|------------|--------------|
| | | /organism="Arachis hypogaea" /db_xref="taxon:3818" | | | |
| | BASE COUNT | 630 a | 473 c | 530 g | 399 t |
| | ORIGIN | | | | |
| | Query Match | 99.4%; | Score 2019.2; | DB 6; | Length 2032; |
| | Best Local Similarity | 99.6%; | Pred. No. 0; | | |
| | Matches 2024; | Conservative | 0; | Mismatches | 8; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| Qy | 1 | aataatcatatatattcatcaatcatctatatataagtagtagcaggagcaaatgagagggag | 60 | | |
| Db | 1 | AATAATCATATATATTCAATCAATCATCTATATATAAGTAGTAGCAGGACCAATGAGAGGGAG | 60 | | |
| Qy | 61 | ggtttctccactgatgctgttgctagggatccttgtccttgctcagtttctgcaacgca | 120 | | |
| Db | 61 | GGTTTCTCCACTGATGCTGTTGCTAGGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGCA | 120 | | |
| Qy | 121 | tgccaagtcatcaccttacccagaagaacaacagagaacccctgcgccagaggtgcctcca | 180 | | |
| Db | 121 | GGCCAAGTCAATCACCTTACCAGAAGAAAACAGAGAACCCCTGCGCCAGAGGTGCTCTCA | 180 | | |
| Qy | 181 | gagttgtcaacaggaacccgatgacttgaagcaaaaggcatgcgagtcctcgtgcaccaa | 240 | | |
| Db | 181 | GAGTTGTCAACAGGAACCCGATGACTTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAA | 240 | | |
| Qy | 241 | gctcgagtatgatcctcgttgtgtctctatgctcgcaggacacacactggcaccaccaacca | 300 | | |
| Db | 241 | GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCGCCACCAACCA | 300 | | |
| Qy | 301 | acgttccctccaggggagcggaacacgtggccgccaacccggagactacgatgatgaccg | 360 | | |
| Db | 301 | ACGTTCCCTCCAGGGAGCGGACACGTTGGCGGCCAACCCGGAGACTACGATGATGACCG | 360 | | |
| Qy | 361 | ccgtcaaccccggaagagaggaaggcccgatggggaccagctggaccgagggagcgtga | 420 | | |
| Db | 361 | CCGTCAACCCCGAAGAGAGGAAGGAGGCCGATGGGGACCAAGCTGAACCGAGGGAGCGTGA | 420 | | |
| Qy | 421 | aagagaagaagactggagacaaccaagaagagattggaggcgaccaagtcacagcagcc | 480 | | |
| Db | 421 | AAGAGAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGTCATCAGCAGCC | 480 | | |
| Qy | 481 | acggaaaaataaggcccgaaaggaagagaaggaagaacaagtggtgggaacaccaggtagcca | 540 | | |
| Db | 481 | ACGGAAAATAAGGCCCGAAGGAAGAGAAGGAAGAACAAAGAGTGGGGAACACCAGGTAGCCA | 540 | | |
| Qy | 541 | tgtgaggggaagaaacatctcggaaacacccctttctacttcccgtcaaggcggttttagcac | 600 | | |
| Db | 541 | TGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGCGGTTTAGCAC | 600 | | |
| Qy | 601 | ccgctacgggaacccaaaacggtaggatccgggtcctgcagaggtttgaccaaagggtcaag | 660 | | |
| Db | 601 | CCGCTACGGGAACCAAAACGGTAGGATCCGGGTCCTGCAGAGGTTTGACCAAAGGTCAAG | 660 | | |
| Qy | 661 | gcagtttcagaaatctccagaatcacccgtatttgcagatcgaggccaaacctaactctc | 720 | | |
| Db | 661 | GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT | 720 | | |
| Qy | 721 | tgttcttcccaagcacgctgatgctgataacatccttgttatccagcaaggggcaagccac | 780 | | |
| Db | 721 | TGTTCTTCCCAAGCACGCTGATGCTGATAACATCCTTGTATTATCCAGCAAGGGCAAGCCAC | 780 | | |
| Qy | 781 | cgtgaccgtagcaaatggcaataacagaaagagctttaacttgcagcagggcccatgcact | 840 | | |
| Db | 781 | CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTAAATCTTGACGAGGGCCATGCAC | 840 | | |
| Qy | 841 | cagaatccccatccgggtttcattctcctacatcttgaaccgccaatgacaaccagaacctcag | 900 | | |
| Db | 841 | CAGAATCCCATCCGGTTTTCATTCTCTACATCTTGAACCGCCATGACAACCCAGAACCTCAG | 900 | | |
| Qy | 901 | agtagctaaaaatctccatgcccttaacacacacccggccagtttgaggatttcttccccgc | 960 | | |
| Db | 901 | AGTAGCTAAAATCTCCATGCCCGGTTAAACACACCCCGCCAGTTTGAGGATTTCTTCCCGGC | 960 | | |

Qy 961 gagcagccgagaccatcatcctacttgagggttcagcaggaatacgttgaggccgc 1020
Db 961 GAGCAGCCGAGACCAATCATCCTACTTGCAGGGATTTCAGCAGGAATACGTTGGAGGCCG 1020
Qy 1021 cttcaatgcggaattcaatgagatacggaggtgctgttagaagagaatgcaggagtga 1080
Db 1021 CTTCAATGCGGAATTCAATGAGATACGGAGGGTGCTGTTAGAAGAGAAATGCAGGAGAGA 1080
Qy 1081 gcaagagagagagggcagagggcgatggagtactcggagttagagaacaatgaagagt 1140
Db 1081 GCAAGAGGAGAGAGGGCAGAGGGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGT 1140
Qy 1141 gatagtcaaagtgtcaaaaggagcacggttgaagaacttactaaagcacgcgtaaatccgtctc 1200
Db 1141 GATAGTCAAAGTGTCAAAGGAGCACGCTTCAAGAACTTACTAAGCACGCTAAATCCGCTCTC 1200
Qy 1201 aaagaaaggctccgaagaagaggagagatatcaccaaccccaatcaacttgagagaagcga 1260
Db 1201 AAAGAAAGGCTCCGAAGAAGAGGGAGATATCACCAACCCCAATCAACTTGAGAGAAGCGGA 1260
Qy 1261 gcccgatctttctaacaactttgggaagtatttggaggtgaagccagacaagaagaaccc 1320
Db 1261 GCCCGATCTTTCTAACAACTTTGGGAGGTTATTTGAGGTGAAGCCAGACACAAGAAACCC 1320
Qy 1321 ccagcttcaggacctggacatgatgctcacctgtgtagagatacaagaaggagcctttgat 1380
Db 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGAGCTTTGAT 1380
Qy 1381 gctcccacacttcaactcaaaaggccatggttatcgtcgtcgtcaacaaaggaaactggaaa 1440
Db 1381 GCTCCCACTTCAACTCAAAAGGCCATGGTTATCGTCTGCTCAACAAAGGAACCTGGAAA 1440
Qy 1441 cettgaactcgtggctgtaagaaaaagagcaacaacagaggggagcggtggaagaagagga 1500
Db 1441 CCTTGAACCTCGTAGCTGTAAGAAAAGAGCAACAACAGAGGGGAGCGGGGGAAGAAGAGGA 1500
Qy 1501 ggaagaaagacgaagaagaggggaagttaacagagaggtgcgttaggtacacagcgaggtt 1560
Db 1501 GGACGAAGACGAAGAAGAGGAGGGGAAGTAACAGAGAGGTTGCGTAGGTACACAGCGAGGTT 1560
Qy 1561 gaaggaaggcgatgtgttcatcatgcccagcagctcatccagtagccatcaacgccttctc 1620
Db 1561 GAAGGAAGCGGATGTGTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
Qy 1621 cgaactccatctgcttggcttcgggtatcaacgctgaaaaacaacacagaatcttccctgc 1680
Db 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCCCTGC 1680
Qy 1681 aggtgataaggacaaatgtgatagaccagatagagaagcaagcgaggtatttagcattccc 1740
Db 1681 AGGTGATAAGGACAAATGTGATAGACCAGATAGAGAAGCAAGCGAAGGATTTAGCATTCCTC 1740
Qy 1741 tgggtcgggtgaacaagttagaagctcatcaaaaaccagaagaatctcactttgtgag 1800
Db 1741 TGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAGAAGGAATCTCACTTTGTGAG 1800
Qy 1801 tgcctcgtcctcaatctcaatctcaatctcgtcgtctcctgagaagaagtcctcctgagaa 1860
Db 1801 TGCTCGTCTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTGAGAAAGAGTCTCCTGAGAA 1860
Qy 1861 agaggaatcaagaggaagaaaaaccaaggaaggaaggtccactcctttcaattttgaagcc 1920
Db 1861 AGAGGATCAAGAGGAGGAGAAACCAAGGAGGAAGGGTCCACTCCTCTTCAATTTGAAGGC 1920
Qy 1921 ttttaactgagaatggaggcaacttggttatgtatcgataataaagatcacgccttttgtact 1980
Db 1921 TTTTAACTGAGAAATGGAGCAACTTGTTATGTATCGATAATAAGATCACGCTTTTGTACT 1980
Qy 1981 ctactatccaaaaacttatcaataaaataaaacggttgtgcgttgtttctcc 2032
Db 1981 CTACTATCCAAAAACTTATCAATAAATAAAAAAGCTTTGTGCGTTGTTCTCTCC 2032

RESULT 4
AX155332
LOCUS AX155332 1978 bp DNA PAT 22-JUN-2001
DEFINITION Sequence 5 from Patent WO0140264.
ACCESSION AX155332
VERSION AX155332.1 GI:14536767
KEYWORDS .
SOURCE peanut.
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 1978)
AUTHORS Bannon,G.A., Burks,W.A., Caplan,M.J., Sampson,H. and Sosin,H.
TITLE Peptide antigens
JOURNAL Patent: WO 0140264-A 5 07-JUN-2001;
Panacea Pharmaceuticals, LLC (US) ; The University of Arkansas (US)
; MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK
(US)
FEATURES
source Location/Qualifiers
1..1978
/organism="Arachis hypogaea"
/db_xref="taxon:3818"
BASE COUNT 614 a 460 c 513 g 391 t
ORIGIN
Query Match 88.8%; Score 1805.2; DB 6; Length 1978;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1935; Conservative 0; Mismatches 43; Indels 54; Gaps 5;
Qy 1 aataatcatatatattcatcatcatctatatataagtagtagcaggagcaatgagaggag 60
Db 1 AATAATCATATATATTTCATCATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGGAG 60
Qy 61 ggtttctccactgatgctgttctagggatccctgtcctggcttcagtttctgcaacgca 120
Db 61 GGTTCCTCCACTGATGCTGTGCTAGGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGCA 120
Qy 121 tgccaaagtcatcaccttaccagaagaacacagagaaacccctgcgccagaggtgcctcca 180
Db 121 GGCCAAAGTCACC-----TTACCGGAAACACAGAGAACCCCTGCGCCAGAGGTGCTCCA 174
Qy 181 gagttgtcaacaggaaaccggatgacttgaagcaaaaagcatcgaggtctcgtgcaccaa 240
Db 175 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAAGGCATGCGAGTCTCGCTGCACCAA 234
Qy 241 gctcagatgatcctcgtgtgtctatgatcctcctcgaggacacactggcaccaccaacca 300
Db 235 GCTCGAGTATGATCCTCGTTGTGTCTATG-----ACACTGGCGCCACCAACCA 282
Qy 301 acgttccccctcaggggagcggaacacgtggccgccaaacccggagactacgatgatgaccg 360
Db 283 ACGTCACCTCCAGGGGAGCGGACACGTCGTGGCCGCAACCCGGAGACTACGATGATGACCG 342
Qy 361 ccgtcaacccccgaagagaggaaggagccgatgggaccagctggaccgagggagcgtga 420
Db 343 CCGTCAACCCCGAAGAGAGAGGAGGCGCCGATGGGGACCAAGCTGAACCCGAGGAGCGTGA 402
Qy 421 aagagaagaagactggagacaacccaagagaagattggaggcgaccaagtcatcagcagcc 480
Db 403 AAGAGAAGAAGACTGGAGACAACCAAGAGAGAAGATTGGAGCGCGACCAAGTCATCAGCAGCC 462
Qy 481 acggaataaaggccccgaaggaagagaagagagaacagtggggaacacaccaggtagcca 540
Db 463 ACGGAAATAAAGGCCCGAAGGAAGAGAAGGAGAACAAAGAGTGGGGAACACCAAGGTAGCGA 522
Qy 541 tgtgaggggaagaaacatctcgggaacaacccctttctacttcccgtcaaggcggttagcac 600
Db 523 GGTGAGGGAAGAAAACATCACGGGAACAACCCCTTTCTACTTCCCCTCAAGGCGGTTAGCAC 582

QY 601 ccgctacgggaaccaaaccggtaggatccgggtcctgcagaggtttgaccaaaaggtcaag 660
|||||
Db 583 CCGCTACGGGAACCAAAACGGTAGGATCCGCGTCTTCGACAGGTTTGACCAAAAGGTCAA 642
QY 661 gcagtttcagaatctccagaatcaccggtattgtgagatcgagggccaaacotaaactct 720
|||||
Db 643 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAGACTA 702
QY 721 tgttcttcccaagcacgctgatgctgataacatccttgttatccagcaagggcaagccac 780
|||||
Db 703 TGTCTTCCCAAGCACGCTGATGCTGTGATAACATCCTTGTATCCAGCAAGACCAAGCCAC 762
QY 781 cgtgaccgtagcaaatggcaataacagaaaagagctttaatcttgacgagggccatgcact 840
|||||
Db 763 CGTGACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAGGCCATGCAC 822
QY 841 cagaatcccatccggtttcatttccctacatcttgaccgccatgacaaaccagaacctcag 900
|||||
Db 823 CAGAAATCCCATCCGTTTTCATTTCTTACATCTTGAAATCGCAATGACAAACCAACCTCAG 882
QY 901 agtagctaaatctccatgcgcgttaacacaccccgccagtttgaggatttcttccccgc 960
|||||
Db 883 AGTAGCTAAAATCTCCATGCGCGTTAACACGCGCCGCGCAGTTTGAGGATTTCTTCCCGGC 942
QY 961 gacgagcagagaccaatcatcctactctgacgttcgagggcttcagcaggaatacgttggagccgc 1020
|||||
Db 943 GAGCAGCGGAGACCAATCATCTACTTTCAGGGGATTCAGCAGGAATACTTTGGAGGCCGC 1002
QY 1021 ctccaatgcggaattcaatgagatacggaggtgctgttagaagagaaatcaggagtg 1080
|||||
Db 1003 CTTCAATCGGGAATTC AATGAGATACGGAGGTGCTGTTAGAAGAGAAATCAGGAGGAGA 1062
QY 1081 gcaagagagagagggcagagggcgatggagtactcggagtagtgagaaacaatgaaggagt 1140
|||||
Db 1063 TCAAGAGAGAGAGGGCAGAGCGGACGGAGTACTCGAGTAGT----TAATGAAGGAGT 1117
QY 1141 gatagcaaaagtgtcaaaaggagcacggttgaaagaacttactaagcacgcgttaaatccgtctc 1200
|||||
Db 1118 GATAGTCAAGTGTCAAAGGAGCACGTTCAAGAACTTACTAAGCACGCTAAATCCGTCTC 1177
QY 1201 aaagaaaggctccgaagaagaggagagatatcaccacccaatcaacttgagagaaaggcga 1260
|||||
Db 1178 AAAGAAAGGCTCCGAAGAAGA--AGATATCACCACCAACCAATCAACTTGAGAGATGGCGA 1234
QY 1261 gccgatctttctaacaactttgggaagtattttgaggtgaagccagacaagaagaaccc 1320
|||||
Db 1235 GCCGATCTTTCTAACAACCTTTGGGAGGTTATTTGAGGTGAAGCCAGACAGAAGAACC 1294
QY 1321 ccagcttcaggaacctggacatgatgctcacctgtgtagagatcaaaagaaggagctttgat 1380
|||||
Db 1295 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAAGAGCTTTGAT 1354
QY 1381 gctcccaacttcaactcaaggcccatggttatactgctgcgtcaacaaaggaaactggaaa 1440
|||||
Db 1355 GCTCCACACTTCAACTCAAAGGCCATGGTTCATCGTCTCGTCAACAAAGGAACTGGAAA 1414
QY 1441 ccttgaactcgtggtgttaagaaaaagagcaacaacagaggggacggcggaagaagagga 1500
|||||
Db 1415 CCTTGAACTCGTAGCTGTAAAGAAAAGAGCAACAACAGAGGGGACGGCGGGAACAAGAGTC 1474
QY 1501 ggacgaagacgaagaagagggggaagtaacagagaggtgcgtaggtacacagcgaggtt 1560
|||||
Db 1475 GGACGAAGACGAAGAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1534
QY 1561 gaaggaggcgatgtgttcacatcatgccagcagctcatccagtagccatcaacgcttcctc 1620
|||||
Db 1535 GAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1594
QY 1621 cgaactccatctgcttggcttcggtatcaacgcgtgaaaaacaaccacagaatcttcttgc 1680
|||||
Db 1595 CGAACTCCATCTGCTTGGCTTCGGGTATCAACGCTGAAAAACAACCACAGAAATCTTCTCTGC 1654
QY 1681 aggtataaggacaatgtgatagaccagatagagaagcaagcgaaggttagcattccc 1740

Db 1655 AGGTGATAAGGACAATGTGATAGACCAGATAGAGAAGCAAGCAAGGATTTAGCATTTCCC 1714
QY 1741 tgggtcgggtgaacaagtgtgagaagctcatcaaaaaaccagaaagaaatctcactttgtgag 1800
|||||
Db 1715 TGGTTCGGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAGTCTCAGTTTGTGAG 1774
QY 1801 tgctcgtcctcaatctcaatctcaatctcgtcgtctcctcgaaaagagctcctcctgagaa 1860
|||||
Db 1775 TGCTCGTCTCAATCTCCGT-----GTCTCCTGAAAA 1806
QY 1861 agaggatcaagaggagaaacccaaggagggaaggtgccactcctttcaattttgaagc 1920
|||||
Db 1807 AGAGGATCAAGAGGAGGAAAAACCAAGGAGGAAGGTCCTCTTCAATTTTGAAGGC 1866
QY 1921 ttttaactgagaatggaggcaacttgttatgtatcgataataagatcacgcttttgtact 1980
|||||
Db 1867 TTTTAACTGAGAAATGGAGGAAACTTGTATGTATCCATAATAAGATCACGCTTTTGTAAAT 1926
QY 1981 ctactatccaaaaacttatcaataataaaaaacgtttgtgcgttgtttctcc 2032
|||||
Db 1927 CTACTATCCAAAACTTATCAATAATAFAAAAACGTTTGTGCGTGTGTTCTCC 1978

RESULT 5
ARQARAH 1949 bp mRNA PLN 10-JAN-1995
LOCUS Arachis hypogea (clone PI7) Ara h I mRNA, complete cds.
DEFINITION Arachis hypogea (clone PI7) Ara h I mRNA, complete cds.
ACCESSION L38853
VERSION L38853.1 GI:620024
KEYWORDS peanut hypersensitivity.
SOURCE Arachis hypogea (strain Florunner) Seed cDNA to mRNA.
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 1949)
AUTHORS Burks,A.W., Cockrell,G., Stanley,J.S., Helm,R.M. and Bannon,G.A.
TITLES Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity
JOURNAL Unpublished
FEATURES Location/Qualifiers
1..1949
/organism="Arachis hypogaea"
/strain="Florunner"
/db_xref="taxon:3818"
/dev_stage="Seed"
1..2
3..1847
/gene="Ara h I"
3..1847
/gene="Ara h I"
/codon_start=1
/protein_id="AAA60336.1"
/db_xref="GI:620025"
/translation="MRGRVSPMLMLLLGILVLASVSATQAKSPYRKTENPCAQRCQLQSC
QEPDDLKQKACESRCTKLEYDPRCVYDTGATNQRHPPGERTGRQPGDYDDRRQPR
REGGRWGPAEPREEREEDWRQPRREDWRPSHQPPRKIRPEGREGEQEWGTPGSEVR
EETSNNPFYFSPRRSTRYGNQGRIRVLQRFQDRSKQFQNLQNHRIQVIEARPNTL
VLPKHADADNLIYIQQOATVTVANGNNRKNRSPNLFNDEGHALRIPSGFISYILNRHDNQ
LRVAKISMPVNTPGQFEFFPASSRDQSSYLOGFSRNTLEAAFNAAFEIRRVLLEEN
AGQEERGORRRSTRSSDNEGVIKVSKEHVQELTKHAKSVSKGSEEDITNPINL
RDGEPDLSNNFGRLEFVKPDKNPQLQDLDMLTCEVEIKEGALMLPHFNSKAMVIVV
NKGTGNLELVAVRKEQQORRREWEDEEEDDEEESGNREVRVRRYARLKEGDFVIMP
AAHPVAINASSELHLLFGINAENNRHIFLAGDXDNVIDQIEKQAKDLAFPGSGEQVE
KLKNQRESHFVSARPOSQSPSPSEKEDQEEENQGGKPLLSILKAFN"
1848..1949
3'UTR
polya_site 599 a 455 c 517 g 378 t
BASE COUNT
ORIGIN

| | | | | | | | | | |
|---|------|---|------|--|--|--|--|--|--|
| Query Match | | | | | | | | | |
| Best Local Similarity 86.2%; Score 1752.4; DB 8; Length 1949; | | | | | | | | | |
| Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6; | | | | | | | | | |
| QY | 48 | caatgagaggaggttctccactgatgctgttgctagggatccttgtcctggttcag | 107 | | | | | | |
| Db | 1 | CAATGAGAGGGAGGGTTCTCCACTGATGCTGTGCTTGGGATCCTTGTCTCGCTTCAG | 60 | | | | | | |
| QY | 108 | tttctgcaacgcatgccaaagtcatcaccttaccagaagaataaacagagaacccctgcgcc | 167 | | | | | | |
| Db | 61 | TTTCTGCAACGCACGCCAAGTCACC-----TTACCGGAAAACAGAGAACCCCTGCGCCC | 114 | | | | | | |
| QY | 168 | agaggtgcctccagattgtcaacaggaaccgggatgacttgaagcaaaaaggcacgcagt | 227 | | | | | | |
| Db | 115 | AGAGGTGCCCTCCAGAGTTGTCAACAGGAACCGGACGACTTGAAGCAAAAAGGCATGCGAGT | 174 | | | | | | |
| QY | 228 | ctcgctgcaccaagctcgagtatgatcctcgttgtgtctatgatcctcgaggacacactg | 287 | | | | | | |
| Db | 175 | CTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATG-----ACACTG | 222 | | | | | | |
| QY | 288 | gcaccaccaaccaacgttccccctccaggggagcggaacacgtgcccgcacaccccgagact | 347 | | | | | | |
| Db | 223 | GCGCCACCAACCAACACGTCACCCTCCAGGGGAGCGGACACGTGGCCGCCAACCCCGGAGACT | 282 | | | | | | |
| QY | 348 | acgatgataccgccgtcaacccccgaagagagaggagggcgatgggaccagctggac | 407 | | | | | | |
| Db | 283 | ACGATGATGCCGCCGCTCAACCCCGAAGAGAGGGAAGGAGGCGGATGGGACCACTGAAC | 342 | | | | | | |
| QY | 408 | cgagggagcgtgaaagagaaagactggagacaaccaagagaagattggaggcgaccaa | 467 | | | | | | |
| Db | 343 | CGAGGGAGCGTGAAAGAGAGAAAGACTGGAGACAACCAAGAGAAGATTGGAGGGGACCAA | 402 | | | | | | |
| QY | 468 | gtcatcagcagccacggaataataaggccccgaaggaagagaagagaacagagtgggaa | 527 | | | | | | |
| Db | 403 | GTCATCAGCAGCCACGGAAATAAAGGCCCGAAGGAAGAGAAGGAGAACAAAGTGGGGAA | 462 | | | | | | |
| QY | 528 | caccaggtagccatgtgagggagaagaacatctcggaaacaccccttctacttcccgta | 587 | | | | | | |
| Db | 463 | CACCAAGTCAAGCAGGTGAGGGRAAGAACATCACGGAACAAACCCTTTCTACTTCCCGTCAA | 522 | | | | | | |
| QY | 588 | ggcgggttagcaccgcgtacgggaaccaaagcgtgtagtccgggtcctgcagaggtttg | 647 | | | | | | |
| Db | 523 | GGCGGTTTAGCACCCGCTACGGGAACCAAACGGTAGGATCCGCGTCTCTGCAGAGGTTG | 582 | | | | | | |
| QY | 648 | accaaaggtcaaggcagtttcagaaatctccagaatcacogtattgtgcagatcgaggcca | 707 | | | | | | |
| Db | 583 | ACCAAAGGTCAAAGCAGTTTCAGAAATCTCCAGAAATCACCGATACCGGTATTGTGCAGATCGAGGCCA | 642 | | | | | | |
| QY | 708 | aacctaacactcttgttcttccaaagcacgctgatgtgtataacatccttgttatccagc | 767 | | | | | | |
| Db | 643 | GACCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGTGATAACATCCTTGTATCCAGC | 702 | | | | | | |
| QY | 768 | aagggaagccaccgtgaccgtagcaaatggcaataacagaaagagcgtttaatcttgacg | 827 | | | | | | |
| Db | 703 | AAGGACAAGCCACCCGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTTAATCTTGACG | 762 | | | | | | |
| QY | 828 | agggccatgcactcagaatcccatccggtttctatttctacatcttgaacccgccaatgaca | 887 | | | | | | |
| Db | 763 | AGGGCCATGCACTCAGAAATCCATCCGGTTTCAATTTCTTACATCTTGAATCGACATGACA | 822 | | | | | | |
| QY | 888 | accagaacctcagagtagctaaaaatctccatgcccggttaacacaccccgccagtttgagg | 947 | | | | | | |
| Db | 823 | ACCAGAACCTCAGAGTAGCTAAATACTCCATGCCCGGTTAACACGCCCGCCAGGTTTGAGG | 882 | | | | | | |
| QY | 948 | atttctccccggcgagcagcgagaccaatcatcctacttgcagggccttcagcaggaata | 1007 | | | | | | |
| Db | 883 | ATTTCTTCCCCGGCAGCAGCCGAGACCAATCATCTACTTGCAGGGATTACGACGGAATA | 942 | | | | | | |
| QY | 1008 | cgttgaggccgccttcaatgcggaattcaatgagatacggaggtgctgttagaagaga | 1067 | | | | | | |
| Db | 943 | CTTTGGAGGCCGCCCTTCAATCGCGGAATTCAATGAGATACGGAGGTGCTGTTAGAAGAGA | 1002 | | | | | | |
| QY | 1068 | atgcaggaggtgagcaagaggagagagggcgagggcgatggagtactcggagtagtgaga | 1127 | | | | | | |

| | | | |
|----|------|--|------|
| Db | 1003 | ATGCAGGAGGAGAGCAAGAGGAGAGAGGGCAGAGCGGACGGAGTACTCGGAGTAGTG--- | 1059 |
| QY | 1128 | acaatgaaggagtgatagtcctcaagtgctcaagagagcacgttgaaagaacttactaagcacg | 1187 |
| Db | 1060 | ATAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCACGTTCAAGAACTTACTAAGCACG | 1119 |
| QY | 1188 | ctaaatccgctctcaagaagagcctccgaagaagagggagatatcaccaacccaatcaact | 1247 |
| Db | 1120 | CTAAATCCGTCTCAAGAGAAAGGCTCC---GAAGAGGAAGATATCACCAACCCCAATCAACT | 1176 |
| QY | 1248 | tgagagaaggcgagcccgatcttcttaacaactttgggaagtatttgaggtgaagccag | 1307 |
| Db | 1177 | TGAGAGATGGCGAGGCCGATCTTTCTAACAACTTTGGGAGGTTATTTGAGGTGAAGCCAG | 1236 |
| QY | 1308 | acaagaagaacccccagcttcaggaccttgacatgatgctcacctgtgtagagatcaag | 1367 |
| Db | 1237 | ACAAGAAAGAACCCCAAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAG | 1296 |
| QY | 1368 | aaggagcttggatgctcccacacttcaactcaaaagccatggttatcgctcgctcaaca | 1427 |
| Db | 1297 | AAGGAGCTTTGATGCTCCACACTTCAACTCAAAAGGCCATGGTTCATCGTCTGTCGTCACACA | 1356 |
| QY | 1428 | aaggaactggaacacttgaactcgtggctgtaagaaaaagagcaacaacagaggggacggc | 1487 |
| Db | 1357 | AAGGAAC TGAAACCTTGAAC TCGTAGCTGTAAGAAAAGAGCAACAACAGAGGGGACGGC | 1416 |
| QY | 1488 | gggaa-----gaagagggaggaaggaaggaaggaaggaaggaaggaaggaaggg | 1538 |
| Db | 1417 | GGGAACAAGAGTGGGAAGAAAGAGGAGGAAGATGAAGAAGAGGAGGGAAGTAACAGAGAGG | 1476 |
| QY | 1539 | tgcgtaggtacacagcgaggttgaaaggaaggcgatgtgttcatcatcgcagcagctcatc | 1598 |
| Db | 1477 | TGCGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATATGCCAGCAGCTCATC | 1536 |
| QY | 1599 | cagtagccatcaacgccttcctccgaactccatctgcttggttcggtatcaacgcctgaaa | 1658 |
| Db | 1537 | CAGTAGCCATCAACGCTTCCTCCGAAC TCCATCTGCTTGGCTTCGGTATCAACGCTGAAA | 1596 |
| QY | 1659 | acaaccacagaatcttccttgaggtgataaaggacaaatgtgatagaccagatagagaagc | 1718 |
| Db | 1597 | ACAACCACAGAACTTTCCTTG CAGGTGATAAGGACAANTGTATAGACCAGATAGAGAAGC | 1656 |
| QY | 1719 | aagcgaaggatttagcatccctgggtcggtggaacaagttgagaagctcatcaaaaaacc | 1778 |
| Db | 1657 | AAGCGAAGGATTTAGCATTCCTCGGTTCGGGTGAACAAGTTGAGAAGCTCATCAAAAAACC | 1716 |
| QY | 1779 | agaaggaatctcaccttgtgagtgctcgctccctcaatctcaatctcccgtcgtctc | 1838 |
| Db | 1717 | AGAGGGAGTCTCACTTTGTGAGTGTCTCGTCCCTCAATCTCAATCTCCGTC----- | 1765 |
| QY | 1839 | ctgagaaagagtcctcctgagaaagaggaatcaagaggaggaaccaaagggaagggtc | 1898 |
| Db | 1766 | -----GTCTCCTGAAAAAGAGGATCAAGAGGAGGAGGAGGAAACCAAGGAGGAGGGTTC | 1815 |
| QY | 1899 | cactcctttcaattttgaaggccttttaactgagaatggaggcaacttgttatgtatcgat | 1958 |
| Db | 1816 | CAC TCTTTCAATTTTGAAGGCTTTTAACTGAGAAATGGAGGAAAACTTGTATGTATCCAT | 1875 |
| QY | 1959 | aataagatcacgcgttttgtacttactatccccaaaaacttatcaataaaaaaacgtttg | 2018 |
| Db | 1876 | AATAAGATCACGCTTTTGTGTAATCTACTATCCAAAAACTTATCAATAAATAAAAAACGTTTG | 1935 |
| QY | 2019 | tgcgttgtttctcc | 2032 |
| Db | 1936 | TGCGTTGTTCTCC | 1949 |

| | | | | |
|------------|---------------------------------|---------|------|-------------|
| RESULT | 6 | | | |
| PSVIC | | | | |
| LOCUS | PSVIC | 1433 bp | mrna | PLN |
| DEFINITION | Pisum sativum mRNA for vicilin. | | | |
| ACCESSION | Y00722 | | | 10-FEB-1999 |


```
REFERENCE 1 (bases 1 to 1955)
AUTHORS Sebastiani,F.L.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1989) Sebastiani F.L., Washington University,
Biology Department, Box 1137, St. Louis, MO 63130, USA
REFERENCE 2 (bases 1 to 1955)
AUTHORS Sebastiani,F.L., Farrell,L.B., Schuler,M.A. and Beachy,R.N.
TITLE Complete sequence of a cDNA of alpha subunit of soybean
beta-conglycinin
JOURNAL Plant Mol. Biol. 15 (1), 197-201 (1990)
MEDLINE 91355860
COMMENT see <J01290> to <J01293> and <J01295> to <J01296> for other Glycine
max seed storage proteins.
FEATURES
source Location/Qualifiers
1..1955
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="pGmc76"
/tissue_type="immature cotyledons"
1..1818
/note="prepro-polypeptide (AA -22 to 583)"
/codon_start=1
/protein_id="CAA35691.1"
/db_xref="GI:18536"
/db_xref="SWISS-PROT:P13916"
/translacion="MMRARFPLLLGLVFLASVSFGLIAYWEKENPKHNKCLQSCNS
ERDSYRNOACHARNLLKVEKEEIGEIPRPQHPEREPQQGSEKEEDEDQPR
PIPFPRPQREHEEQEWEPRKEKEGSEEEDEDEDEQDFQFPFRPPH
QKEERNEEDEDQRESESESDSELRRHKNKPNFLFGSNRFETLFKNQYGRIRVLQ
RFNQRSPOLNLDYRILFENSKPNTLLPNHADADYLIIVLNGTALISLVNDDRDS
YRLQSGDALRVPSTGYVVPDNNENLRITLAI PVNKPGRFESFFLSSTEAQQSYL
QGFSRNILEASYDTKFEINKVLFREEQQQGQRLQESVIVEISKEQIRALSKRK
SSSRKTISSDKPFLNRSDPIYSNKLKGFETPEKNPQLRDLDFLSIVDMNEGAL
LLPHFNSKAIVILVINEGDANIELVGLKEQQEQEQLPVRKYRAELSEQDIFVI
PAGYPVVVNATSNLNFPAIGINAENNORNFLAGSDQNVISQIPSOVQELAFPGSAQAV
EKLLKNQRESYFVDAQPKKKEGNGRKGKPLSLIRAFY"
sig_peptide 1..66
misc_feature 67..1815
mat_peptide 187..1815
old_sequence 594..597
polya_site 1955
BASE COUNT 625 a 444 C 455 g 431 t
ORIGIN
Query Match 19.9%; Score 404; DB 8; Length 1955;
Best Local Similarity 61.2%; Pred. No. 4.8e-101;
Matches 765; Conservative 0; Mismatches 410; Indels 75; Gaps 4;
QY 563 aacaaccccttctacttcccgtaagcggttagcacccgctacgggaacaaaacggt 622
Db 577 AAGAACCCCTTTTCTCTTCGGCTCTAACAGGTTTCGAAACTCTCTTCAAAAACCAATATGGT 636
QY 623 aggatccgggtcctgcagaggtttgacaaaaggtaaggcagtttcagaatctccagaat 682
Db 637 CGCATTCGGTCTCTCCAGAGGTTCAACCAACGCTCCCCACAACTTCAGAACTCTCCGAGAC 696
QY 683 caccgtattgtgcagatcgaggccaaaacctaaactcttgttcttcccaagcacgctgat 742
Db 697 TACCGCATTTTGGAGTTCAACTCCAAACCCCAACACCTCTTCTCCCCAACCATGCTGAC 756
QY 743 gctgataacatccttggttatccagcaagggcaagcccgtagccgtagcaaatggcaat 802
Db 757 GCTGATTACCTCATCGTTATCCTTAACGGGACTGCCATCTTTTCTTGGTGAACAACGAC 816
QY 803 aacagaaagagctttaatcttgacaggggccatgcactcagaatcccatccggtttcatt 862
Db 817 GACAGAGACTCCTACAGACTTCAATCTGGTGATGCCCTTGAGAGTGCCCTCAGGAACACACA 876
```

```
QY 863 tcctacatcttgaaccgcccatgataaccagaaacctcagagtagctaaaaatctccatgccc 922
Db 877 TACTATGTGGTCAACCCCTGACAACAACGAAATCTCAGATTAAATAACACTCGCCATACC 936
QY 923 gttaacacacccggccagtttgaggatttcttcccgcgagcagcgagaccaatcatcc 982
Db 937 GTTAACAAGCCTGGTAGATTGAGAGTTTCTTCTATCTAGCACTGAAGCTCAACAATCC 996
QY 983 tacttgacgggttcacgaggaatacgttggaggccgccttcaatgcggaattcaatgag 1042
Db 997 TACTTGCAAGGATTACGACGGAACATTTAGAGGCCCTCCTACGATACCAAAATTCGAGGAG 1056
QY 1043 atacggaggtgtgttagaagaaatgcaggaggtgagcaagaggagagaggcgagag 1102
Db 1057 ATAAACAAGGTTCTGTTTAGTAGAGGAAGGCGACGCAAGGGGACAGAGGCTG--- 1113
QY 1103 cgatggagtactcggagttagtgagaacaatgaaggagtgtatgtcaaaagtgtcaaaaggag 1162
Db 1114 -----CAAGAGAGCGGTGATTGTGGAAATCTCGAAGGAA 1146
QY 1163 cactgtgaagaacttactaagcacgcgtaaatccgtctcaaaagaaaggctccgaagaagag 1222
Db 1147 CAGATTCGGGCACTGAGCAAAACGTCGCAAAATCTAGTTCAAGGAAACCATTTCTTCTGAA 1206
QY 1223 ggagatatcaccaccccaatcaacttgagagaaggcgagccccgatcttcttaacaacttt 1282
Db 1207 G-----ATAAACCTTTAACTTGAGAAAGCCGCGACCCCATCTACTCCAACAAGCTT 1257
QY 1283 gggaaagttatttgagtgaaagccagacaaagaagacccccagcttccagacctggacatg 1342
Db 1258 GGCAAGTCTTTGAGATCACCCCG--AGAAAAACCCCGAGCTTCGGGACTTGGATATC 1314
QY 1343 atgctcacctgtgtagagatcaaaagaaggagctttagtgcctcccacacttcaactcaaag 1402
Db 1315 TTCTCTAGTATTGTGGATATGAACGAGGGAGCTCTTCTTCTACCCACACTCAATTCAAAG 1374
QY 1403 gccatggttatcgtcgtctcaacaaaggaaactggaaaccttgaactcgtggtgtaaga 1462
Db 1375 GCGATAGTGATACTGGTAATTAAATGAAGGAGATGCAAAACATTGAACTTGTGGCCTAAA 1434
QY 1463 aaagagcaacaacagaggggacggcggtacacagcgaggttgaaggaaagcgatgtgttcac 1522
Db 1435 GAACAACAACAGGAGCAGCAACAGGAAGAGACCTTTG----- 1473
QY 1523 ggaagtaacagagaggtgcgttaggtacacagcgaggttgaaggaaagcgatgtgttcac 1582
Db 1474 -----GAAGTGGGAAATATAGAGCCGAATTGTCTGAACAAGATATATTGTA 1521
QY 1583 atgccagcagctcatccagtagccatcaacgcgttctccgaaactccatcctgctggttc 1642
Db 1522 ATCCAGCAGGTTATCCAGTTGTGGTCAACGCTACCTCAAATCTGAATTTCTTGTCTATT 1581
QY 1643 ggtataacgctgaaaaaacacacagaatcttcttgcaggtgataaaggacaatgtgata 1702
Db 1582 GGTATTAAATGCCGAGAACACACAGAGGAACTTCCTCGCAGGTTCCGCAAGACATGTGATA 1641
QY 1703 gaccagatagagaagcaagcgaaggatttagcattccctcggtcggtggaacaagttgag 1762
Db 1642 AGCCAGATACCTAGTCAAGTGCAGGAGCTTGCAATTCCTGGGTCTGCACAAGCTGTGTGAG 1701
QY 1763 aagctcatcaaaaaaccagagaatctcactttgtgagtgctgcctcctca 1812
Db 1702 AAGCTATTAAAGAAACCAAGAGAAATCCTACTTTTGTGGATGCTCAGCCTAA 1751
RESULT 13
AB008680
LOCUS AB008680 1680 bp mRNA PLN 01-SEP-2000
DEFINITION Glycine max mRNA for alpha' subunit of beta-conglycinin, partial
cds.
ACCESSION AB008680
VERSION AB008680.2 GI:9967360
```


KEYWORDS alpha' subunit of beta-conglycinin.
SOURCE Glycine max (cultivar:Wasesuzunari) cDNA to mRNA.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (sites)
AUTHORS Maruyama,N., Katsube,T., Wada,Y., Oh,M.H., Barba De La Rosa,A.P., Okuda,E., Nakagawa,S. and Utsumi,S.
TITLE The roles of the N-linked glycans and extension regions of soybean beta-conglycinin in folding, assembly and structural features
JOURNAL Eur. J. Biochem. 258 (2), 854-862 (1998)
MEDLINE 99089652
REFERENCE 2 (bases 1 to 1680)
AUTHORS Maruyama,N., Katsube,T., Wada,Y. and Utsumi,S.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1997) Shigeru Utsumi, Kyoto University, Research Institute for Food Science; Gokanoshio, Uji, Kyoto 611, Japan (E-mail:utsumi@soya.food.kyoto-u.ac.jp, Tel:0774-38-3760, Fax:0774-38-3761)
COMMENT On Sep 5, 2000 this sequence version replaced gi:4191813.
FEATURES
source Location/Qualifiers
1..1680
/organism="Glycine max"
/cultivar="Wasesuzunari"
/db_xref="taxon:3847"
CDS
<1..1680
/codon_start=1
/product="alpha' subunit of beta-conglycinin"
/protein_id="BA74452.2"
/db_xref="GI:9967361"
/translation="VEEEEEEGQIPRRPQHPRERQKHGEKEDEGEQRPFPFP
RPRPQHEEHEQEEHWRKKEKHGGKSEEEQDEREHPRPQPHQKEEKEHWOH
KQKHQKSESEEEEDODEQKESQESQREPRRHKNKNPFHNSKRFQTL
FNQYGHVRLQRFNRSQQLNLRDYRILEFNSKPNLTLPHZADADYLIVLNGTA
ILTLVNNDDSDSYNLOSQDALRVPAQTTYYVVPNDENLRLMITLAIPVKNKGRFESF
FLSTQAQSYLQFSKNILEASYDTKFEINKVLFGREGQQQGEERLQESVIVEIS
KQIRELSKHAKSSRKTISSEDKPFLNRSRDP IYSNKLGLKFEITPEKNPOLRLDV
FLSVVDMNEGALFLPHNSKAIIVLVINEGEANIELVGIKEQQQRQQEEQPLEVRKY
RAELSEQDIFVIPAGYPVVVNATSDLNFFAFGINAENNRNFLAGSKDNVISQIPSOV
QELAFPGSAKDNIENLKSQSESYFVDAQPQKQKEGNGKRGKGPLSLIRAFY"
BASE COUNT 567 a 385 c 395 g 333 t
ORIGIN
Query Match 19.8%; Score 402; DB 8; Length 1680;
Best Local Similarity 61.6%; Pred. No. 1.7e-100;
Matches 770; Conservative 0; Mismatches 405; Indels 75; Gaps 5;
QY 563 acaaaccttctactccgtcaaggcggttagcaccccgctacgggaacccaaacgggt 622
|| ||||| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
Db 439 AAGAACCTTTTCACTCAACTCTAAAGGTTCCAAACTCTCTTCAAAAACCAATATGGC 498
QY 623 aggatccgggtcctgcagaggtttgaccaaagggtcaaggcagtttcagaaatctccagaat 682
|| ||||| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
Db 499 CACGTCGCGTCTCCAGAGGTTCAACAAACACGCTCCCAACACAGCTTCAGAAATCTCCGAGAC 558
QY 683 caccgtattgtgcagatcgaggccaaacctaacactcttgtttcttcccaagcacgctgat 742
|| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 559 TACCGCATTTTGGAGTTCAACTCCAAACCCAAACACCCCTTCTTCTCCCCCACCATGCTGAC 618
QY 743 gctgataacatccttgttatccagcaaggcggaagccaccgtagcagcaaatggcaat 802
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 619 GCTGATTACCTCATCTGTTATCCTTAACGGGACTGCCATTCTTACCTTGGTGAACACGAC 678
QY 803 aacagaaagagctttaatcttgcagcagggcgccatgcactcagaatccccatccggtttcatt 862
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 679 GACCGAGACTCTTACAACTTCAATCTGGCGATGCCCTAAGAGTCCCTGCAGGAACACA 738
QY 863 tctacatcttgaacggccatgacaaaccagaacctcagagtagctaaaaatctccatgcc 922
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 739 TACTATGTGGTTAACCCCTGACAACGACGAGAATCTCAGAATGATAACACTCGCCATACCC 798

QY 923 gttaacacacccggccagtttgaggatttcttcccgcgagcagcagcagaccaatcatcc 982
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 799 GTTAACAAACCCGGTAGATTGAGAGTTTCTTCCTATCTAGCACCTCAAGCTCAACAGTCC 858
QY 983 tacttgagggtcttcagcaggaatacgttggagccgccttcaatgcggaattcaatgag 1042
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 859 TACTTGCAAGGGTTCAGCAAGAAATATTCTAGAGGCTCATACGACACCAAAATTCGAGGAG 918
QY 1043 atacggagggtgctgtttaagaagaaatgcaggaggtgagcaagagagagggcagagg 1102
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 919 ATAAACAAGGTTCTGTGTTGTAGAGAGGAGGGGCGACAAACAAGGGGAGAGAGGCTG-- 975
QY 1103 cgatggagtactcgggagttagtgagaacaaatgaaggagttagtcaaaagtgtcaaaaggag 1162
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 976 -----CAAGAGAGTGTGATTGTGGAATCTCAAAGAAA 1008
QY 1163 cacgttgaagaacttactaagcacgctaaatccgtctcaaaagaaaggctccgaaagag 1222
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1009 CAAATTCGGGAACCTGAGCAAAACATGCCAAATCTAGTTCAAGGAAACCATTTCTTCTGAA 1068
QY 1223 ggagatatcaccaaccaatcaacttgagagaaggcgagccccgatcttcttaacaaacttt 1282
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1069 G-----ATAACCTTTCAACTTGAGAAAGCCGACCCCATCTATTCCAACAAGCTT 1119
QY 1283 ggaagtattttaggtgaagccagacaagaagaacccccagcttcaggacctggacatg 1342
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1120 GGCAAGTTGTTGAGATTACCCAG--AGAAAAACCCCTCAGCTTCGGGACTTGGATGTC 1176
QY 1343 atgctcacctgtgtagagatcaaaagaaggagcttttgatgctccccacacttcaactcaaaag 1402
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1177 TTCTCAGTGTGTGGATATGAACGAGGGAGCTCTTTTCTGCCACACTTCAATTCAAAG 1236
QY 1403 gccatggttatcgtcgtcaacaaaaggaaactggaacaccttgaactcgtggctgtaaga 1462
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1237 GCCATAGTGGTACTAGTGATTAAATGAAGGAGGAAGCAACATTGAACCTGT---TGGCATT 1293
QY 1463 aaagacaaacaacagaggggagcggggaagaagaaggagggagacgaagaagaagagag 1522
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1294 AAAGAACAACAACAGAGGCGCACAGGAAGAGCAACCTTTGGAA----- 1338
QY 1523 ggaagtaacagagaggtgcgtaggttacacagcgaggttgaaaggaaggcgatgtgttcac 1582
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1339 -----GTGCGGAATATAGAGCTGAATTGTCTGAACAAGATATATTGTA 1383
QY 1583 atgccagcagctcatccagtagccatcaacgcgttctccggaactccatctgcttggttc 1642
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1384 ATCCCAGCAGGTTATCCAGTTGTGGTCAACGCTACCTCAGATCTGAATTTCTTTGCTTTT 1443
QY 1643 ggtataacgctgaaaaaacacacagaaatcttcttcaggttgataaagacaatgtgata 1702
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1444 GGTATCAATGCCGAGAACACCAGAGGAACTTCCTTCGAGGTTTCGAAAGACAATGTGATA 1503
QY 1703 gaccagatagagaagcaagcagaggatttagcattccctgggtcgggtgacaagttag 1762
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1504 AGCCAGATACCTAGTCAAGTGCAGGAGCTTGCCTTCCCTGGGTCTGCAAAAAGATATTGAG 1563
QY 1763 aagctcatcaaaaaaccagaaagtatctcacttttgcagtgctcgtcctca 1812
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
Db 1564 AACCTAATAAAGAGCCAAAGTAGTCTACTTTGTGATGCTCAGCCTCA 1613

RESULT 14
PSA276875
LOCUS PSA276875 2069 bp mRNA PLN 27-MAR-2000
DEFINITION Pisum sativum mRNA for convicilin (cvc gene).
ACCESSION AJ276875
VERSION AJ276875.1 GI:7339550
KEYWORDS convicilin; cvc gene; seed storage protein.
SOURCE pea.
ORGANISM Pisum sativum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 13:15:28 ; Search time 486.97 Seconds
(without alignments)
3577.397 Million cell updates/sec

Title: US-09-715-036-5
Perfect score: 2032
Sequence: 1 aataatcatatatattcatc.....cgtttgcggtgtttctcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

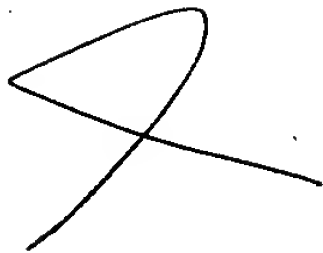
SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 2032 | 100.0 | 2032 | 22 | AAF90339 |
| 2 | 2032 | 100.0 | 2041 | 22 | AAAS08537 |
| 3 | 2025.6 | 99.7 | 2032 | 18 | AAT76613 |
| 4 | 1930 | 95.0 | 1930 | 20 | AAZ06382 |
| 5 | 1752.4 | 86.2 | 1949 | 18 | AAT76612 |
| 6 | 1749.8 | 86.1 | 1952 | 22 | AAAS08539 |
| 7 | 406.8 | 20.0 | 1320 | 19 | AAV17564 |
| 8 | 404 | 19.9 | 1818 | 19 | AAV17562 |
| 9 | 324.2 | 16.0 | 1920 | 19 | AAV17563 |
| 10 | 160.2 | 7.9 | 2140 | 19 | AAV42316 |
| 11 | 157.4 | 7.7 | 2171 | 19 | AAV42311 |
| | | | | | Peanut allergen Ar |
| | | | | | DNA encoding anaph |
| | | | | | Peanut allergen Ar |
| | | | | | Peanut allergen, A |
| | | | | | Peanut allergen Ar |
| | | | | | Anaphylactic antigen |
| | | | | | Coding sequence fo |
| | | | | | Coding sequence fo |
| | | | | | Coding sequence fo |
| | | | | | Macadamia integrif |
| | | | | | Macadamia integrif |

| | | | | | | |
|------|-------|-----|-------|----|----------|--------------------|
| 12 | 152.6 | 7.5 | 2171 | 19 | AAV42310 | Macadamia integrif |
| 13 | 138.8 | 6.8 | 1867 | 13 | AAQ20377 | Sequence of 67 kD |
| 14 | 98.8 | 4.9 | 1924 | 20 | AAV72243 | G. max SBP2 cDNA. |
| C 15 | 53.2 | 2.6 | 936 | 22 | AAF58252 | Oligonucleotide D1 |
| C 16 | 53.2 | 2.6 | 936 | 22 | AAF58252 | Oligonucleotide D1 |
| C 17 | 53.2 | 2.6 | 936 | 22 | AAF58257 | Oligonucleotide D1 |
| C 18 | 53.2 | 2.6 | 936 | 22 | AAF58259 | Oligonucleotide D2 |
| C 19 | 53.2 | 2.6 | 936 | 22 | AAF58262 | Oligonucleotide D2 |
| C 20 | 53.2 | 2.6 | 938 | 22 | AAF58255 | Oligonucleotide D1 |
| 21 | 49.6 | 2.4 | 10732 | 21 | AAA10594 | Gene encoding a su |
| 22 | 49.4 | 2.4 | 1068 | 22 | AAD05672 | Human zmsel cDNA. |
| 23 | 47.6 | 2.3 | 936 | 22 | AAF58252 | Oligonucleotide D1 |
| 24 | 47.6 | 2.3 | 936 | 22 | AAF58254 | Oligonucleotide D1 |
| 25 | 47.6 | 2.3 | 936 | 22 | AAF58257 | Oligonucleotide D1 |
| 26 | 47.6 | 2.3 | 936 | 22 | AAF58259 | Oligonucleotide D2 |
| 27 | 47.6 | 2.3 | 936 | 22 | AAF58262 | Oligonucleotide D2 |
| 28 | 47.6 | 2.3 | 938 | 22 | AAF58255 | Oligonucleotide D1 |
| C 29 | 45.2 | 2.2 | 1823 | 21 | AAC35016 | Arabidopsis thalia |
| C 30 | 45.2 | 2.2 | 1823 | 21 | AAC36549 | Arabidopsis thalia |
| 31 | 44 | 2.2 | 4000 | 18 | AAT91902 | Mannose-1-phosphat |
| 32 | 43.8 | 2.2 | 441 | 21 | AAA45042 | Mouse secreted exp |
| 33 | 42.4 | 2.1 | 726 | 22 | AAI22214 | Probe #12147 for g |
| 34 | 42.4 | 2.1 | 726 | 22 | AAI47511 | Probe #16197 used |
| 35 | 42.4 | 2.1 | 726 | 22 | AAI07913 | Probe #7904 used t |
| 36 | 42.4 | 2.1 | 1951 | 22 | AAI12976 | Probe #2909 for ge |
| 37 | 42.4 | 2.1 | 1951 | 22 | AAI34335 | Probe #3021 used t |
| 38 | 42.4 | 2.1 | 1951 | 22 | AAI02894 | Probe #2885 used t |
| C 39 | 41.8 | 2.1 | 227 | 22 | AAI22239 | Probe #12172 for g |
| C 40 | 41.8 | 2.1 | 227 | 22 | AAI47532 | Probe #16218 used |
| C 41 | 41.8 | 2.1 | 227 | 22 | AAI07937 | Probe #7928 used t |
| C 42 | 41.8 | 2.1 | 494 | 22 | AAI13001 | Probe #2934 for ge |
| C 43 | 41.8 | 2.1 | 494 | 22 | AAI34356 | Probe #3042 used t |
| C 44 | 41.8 | 2.1 | 494 | 22 | AAI02918 | Probe #2909 used t |
| 45 | 41.2 | 2.0 | 575 | 22 | AAI23315 | Probe #13248 for g |

ALIGNMENTS

| | |
|----------|--|
| RESULT 1 | |
| AAF90339 | |
| ID | AAF90339 standard; cDNA; 2032 BP. |
| XX | |
| AC | AAF90339; |
| XX | |
| DT | 23-JUL-2001 (first entry) |
| XX | |
| DE | Peanut allergen Ara h1 P41B cDNA. |
| XX | |
| KW | Peanut; allergen; Ara h1 P41B; transgenic plant; allergy; ss. |
| XX | |
| OS | Arachis hypogea. |
| XX | |
| FH | Key |
| FT | misc_feature |
| FT | 676..930 |
| FT | /*tag= a |
| FT | /note= "PCR amplified fragment" |
| XX | |
| PN | WO200136621-A2. |
| XX | |
| PD | 25-MAY-2001. |
| XX | |
| PF | 20-NOV-2000; 2000WO-US31657. |
| XX | |
| PR | 19-NOV-1999; 99US-0167255. |
| XX | |
| PA | (UYAL-) UNIV ALABAMA A & M. |
| XX | |
| PI | Dodo HW, Arntzen CJ, Konan KN, Viquez OM; |
| XX | |
| DR | WPI; 2001-355630/37. |
| XX | |
| PT | Producing transgenic peanut plants that produce allergen-free seeds, |



PT useful in non-allergenic foods, by antisense or sense co-suppression of
PT allergen-encoding genes -

XX
PS Claim 20; Fig 5; 72pp; English.

XX
CC The present sequence is that of peanut allergen Ara h1 P41B cDNA.
CC A portion of this gene is homologous to the corresponding region
CC of the peanut allergen Ara h1 P17 gene. This region has been PCR
CC amplified, cloned in transformation vectors (pUC18 and pBI4134) in
CC sense and antisense orientations and used to down-regulate Ara h1
CC P41B and Ara h1 P17 allergens in peanut. This is an example of the
CC method of the invention, which relates to the production of a
CC peanut plant having reduced, or undetectable, allergenic protein
CC (AP) content in its seed. A peanut plant cell is transformed with
CC a DNA construct containing an antisense AP gene and/or sense AP
CC gene, or their fragments, regenerated to plants, and fertile
CC transgenic plants that produce seeds with reduced AP content are
CC identified. The seeds are useful for preparation of allergen-free
CC foods.

XX
SQ Sequence 2032 BP; 628 A; 473 C; 530 G; 401 T; 0 other;

Query Match 100.0%; Score 2032; DB 22; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | aataatcatatatattcatcaatcatcatatataaagtagtagcaggagcaatgagagggag | 60 |
| Db | 1 | aataatcatatatattcatcaatcatcatatataaagtagtagcaggagcaatgagagggag | 60 |
| QY | 61 | ggtttctccactgatgctgttgctagggatccttgtcctggttcagtttctgcaacgca | 120 |
| Db | 61 | ggtttctccactgatgctgttgctagggatccttgtcctggttcagtttctgcaacgca | 120 |
| QY | 121 | tgccaagtcatcaccttaccagaagaacacagagaaacccctgcgccagaggtgcctcca | 180 |
| Db | 121 | tgccaagtcatcaccttaccagaagaacacagagaaacccctgcgccagaggtgcctcca | 180 |
| QY | 181 | gagttgtcaacaggaaacccggatgactgaagcaaaaaggcaatgcgagtctcgctcacc | 240 |
| Db | 181 | gagttgtcaacaggaaacccggatgactgaagcaaaaaggcaatgcgagtctcgctcacc | 240 |
| QY | 241 | gctcgagtagtgcctcctgctgtgtctatgatcctcagaggaacacactggcaccaca | 300 |
| Db | 241 | gctcgagtagtgcctcctgctgtgtctatgatcctcagaggaacacactggcaccaca | 300 |
| QY | 301 | acgttccccccaggggagcggacacagtggcgcccaacccggagactacgatgaccg | 360 |
| Db | 301 | acgttccccccaggggagcggacacagtggcgcccaacccggagactacgatgaccg | 360 |
| QY | 361 | ccgtcaacccccgaagagaggaagggccgatggggaccagctggaccgagggagcgtga | 420 |
| Db | 361 | ccgtcaacccccgaagagaggaagggccgatggggaccagctggaccgagggagcgtga | 420 |
| QY | 421 | aagagaagaagactggagacaacccaagagaagattggaggcgaccaagtcatcagcagcc | 480 |
| Db | 421 | aagagaagaagactggagacaacccaagagaagattggaggcgaccaagtcatcagcagcc | 480 |
| QY | 481 | acggaaaaataagccccgaaggaaagagaagagaacaagagtggggaaacaccaggtagcca | 540 |
| Db | 481 | acggaaaaataagccccgaaggaaagagaagagaacaagagtggggaaacaccaggtagcca | 540 |
| QY | 541 | tgtgaggggaagaacatctcggaaacacccctttctacttcccgtcaaggcggttagcac | 600 |
| Db | 541 | tgtgaggggaagaacatctcggaaacacccctttctacttcccgtcaaggcggttagcac | 600 |
| QY | 601 | ccgctacggggaacaaaaacggtagtatccgggtcctgcagaggtttgacccaaaggtcaag | 660 |
| Db | 601 | ccgctacggggaacaaaaacggtagtatccgggtcctgcagaggtttgacccaaaggtcaag | 660 |
| QY | 661 | gcagtttcagaatatctccagaatcacccgtattgtgcagatgcagggccaaacctaactct | 720 |
| | | | |

| | | | |
|----|------|--|------|
| Db | 661 | gcagtttcagaatatctccagaatcacccgtattgtgcagatcgaggccaaacctaacactct | 720 |
| QY | 721 | tgttcttcccaagcacgctgatgctgataaacatccttgttatccagcaagggcaagccac | 780 |
| Db | 721 | tgttcttcccaagcacgctgatgctgataaacatccttgttatccagcaagggcaagccac | 780 |
| QY | 781 | cgtgaccgtagcaaaaatggcaataaacagaaaagcgtttaatcttgcagaggccatgcact | 840 |
| Db | 781 | cgtgaccgtagcaaaaatggcaataaacagaaaagcgtttaatcttgcagaggccatgcact | 840 |
| QY | 841 | cagaatccccatccggtttcatttcctacatcttgaaccgcccataacaccagaacctcag | 900 |
| Db | 841 | cagaatccccatccggtttcatttcctacatcttgaaccgcccataacaccagaacctcag | 900 |
| QY | 901 | agtagctaaaaatctccatgcccggttaaacacaccccgccagtttgaggatttcttccggc | 960 |
| Db | 901 | agtagctaaaaatctccatgcccggttaaacacaccccgccagtttgaggatttcttccggc | 960 |
| QY | 961 | gagcagccgagaccaatcatcctacttgcagggttcagcagggaatacgttggaggccgc | 1020 |
| Db | 961 | gagcagccgagaccaatcatcctacttgcagggttcagcagggaatacgttggaggccgc | 1020 |
| QY | 1021 | cttcaatgcgggaattcaatgagatacggagggtgctgttagaagaagaatgcaggaggtga | 1080 |
| Db | 1021 | cttcaatgcgggaattcaatgagatacggagggtgctgttagaagaagaatgcaggaggtga | 1080 |
| QY | 1081 | gcaagaggagagagggcgagggcgatggagtactcggagtagtgagaacaatgaaggagt | 1140 |
| Db | 1081 | gcaagaggagagagggcgagggcgatggagtactcggagtagtgagaacaatgaaggagt | 1140 |
| QY | 1141 | gatagtcaaaagtgtcaaaaggagcacgttgaaagaacttactaagcacgctaaaatccgtctc | 1200 |
| Db | 1141 | gatagtcaaaagtgtcaaaaggagcacgttgaaagaacttactaagcacgctaaaatccgtctc | 1200 |
| QY | 1201 | aaagaaaggctccgaagagaggagagatatcaccaaccccaatcaacttgagagaagcgga | 1260 |
| Db | 1201 | aaagaaaggctccgaagagaggagagatatcaccaaccccaatcaacttgagagaagcgga | 1260 |
| QY | 1261 | gccgatcttctaaacaactttgggaagttattggggtgaagccagacaagaagaaccc | 1320 |
| Db | 1261 | gccgatcttctaaacaactttgggaagttattggggtgaagccagacaagaagaaccc | 1320 |
| QY | 1321 | ccagcttcaggacctggacatgatgctcacctgtgtagagatcaaaagaaggagctttgat | 1380 |
| Db | 1321 | ccagcttcaggacctggacatgatgctcacctgtgtagagatcaaaagaaggagctttgat | 1380 |
| QY | 1381 | gctccacacttcaactcaaaagccatggttatcgtcgtcaacaaaaggaaactggaaa | 1440 |
| Db | 1381 | gctccacacttcaactcaaaagccatggttatcgtcgtcgtcaacaaaaggaaactggaaa | 1440 |
| QY | 1441 | ccttgaactcgtggctgtaagaaaaagagcaacaacagaggggacgcgggaaagaagga | 1500 |
| Db | 1441 | ccttgaactcgtggctgtaagaaaaagagcaacaacagaggggacgcgggaaagaagga | 1500 |
| QY | 1501 | ggacgaagacgaagaagggagggaagtaaacagagaggtgcgtaggtacacagcgaggtt | 1560 |
| Db | 1501 | ggacgaagacgaagaagggagggaagtaaacagagaggtgcgtaggtacacagcgaggtt | 1560 |
| QY | 1561 | gaaggaaaggcgatgtgttcatcatgccagcgctcatccagtagccatcaacgcttccctc | 1620 |
| Db | 1561 | gaaggaaaggcgatgtgttcatcatgccagcgctcatccagtagccatcaacgcttccctc | 1620 |
| QY | 1621 | cgaactccatctgcttgcttcgggtatcaacgctgaaaaacaacacagaaatcttccctgc | 1680 |
| Db | 1621 | cgaactccatctgcttgcttcgggtatcaacgctgaaaaacaacacagaaatcttccctgc | 1680 |
| QY | 1681 | aggtgataaggacaatgtgatagaccagatagagaagcaagcgaggatttagcattccc | 1740 |
| Db | 1681 | aggtgataaggacaatgtgatagaccagatagagaagcaagcgaggatttagcattccc | 1740 |
| QY | 1741 | tgggtcggggtgaacaaagttagaaagctcatcaaaaaaccagaagaatctcactttgtgag | 1800 |
| Db | 1741 | tgggtcggggtgaacaaagttagaaagctcatcaaaaaaccagaagaatctcactttgtgag | 1800 |

QY 1801 tgctgcctcaatctcaatctcaatctcgcgtcgtctcctgagaaagagtgctcctgagaa 1860
|||||
Db 1801 tgctgcctcaatctcaatctcaatctcgcgtcgtctcctgagaaagagtgctcctgagaa 1860
|||||
QY 1861 agaggatcaagagaggagaaacccaagaggagggtccactccttcaattttgaaggc 1920
|||||
Db 1861 agaggatcaagagaggagaaacccaagaggagggtccactccttcaattttgaaggc 1920
|||||
QY 1921 ttttaactgagaatggaggcaacttgttatgtatcgataataagatcacgcttttgtact 1980
|||||
Db 1921 ttttaactgagaatggaggcaacttgttatgtatcgataataagatcacgcttttgtact 1980
|||||
QY 1981 ctactatccaaaacttatcaataaaataaaaacggttgtgcgttgtttctcc 2032
|||||
Db 1981 ctactatccaaaacttatcaataaaataaaaacggttgtgcgttgtttctcc 2032
|||||

RESULT 2
AAS08537
ID AAS08537 standard; cDNA; 2041 BP.
XX
AC AAS08537;
XX
DT 23-OCT-2001 (first entry)
XX
DE DNA encoding anaphylactic antigen Ara h 1.
XX
KW Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 50..1930
FT /*tag= a
FT /transl_except= (pos:902..904, aa:Ala)
FT /product= "Peptide antigen Ara h 1"

WO200140264-A2.
07-JUN-2001.
06-DEC-2000; 2000WO-US33124.
06-DEC-1999; 99US-0455294.
23-JUN-2000; 2000US-0213765.
27-SEP-2000; 2000US-0235797.
(PANA-) PANACEA PHARM LLC.
(UYAR-) UNIV ARKANSAS.
(MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
WPI; 2001-381378/40.
P-PSDB; AAU04706.

Antigenic fragments useful for reducing anaphylactic risk and reducing the severity and/or number of allergic symptoms in individuals sensitive to antigens, have reduced ability to bind Immunoglobulin E -
Disclosure; Fig 9; 100pp; English.
The sequence represents the coding sequence of anaphylactic antigen Ara h 1. Ara h 1 is an anaphylactic antigen (A), which was used to design antigenic peptides having a reduced ability to bind immunoglobulin E (IgE) as compared with the intact (A), or having a sequence substantially identical to a portion of sequence of an antigen that includes at least one IgE binding site, where at least one IgE binding site of the peptide is altered. The antigenic peptides are used in a composition which is useful for reducing risk or severity of allergic reaction to an antigen. This is done by identifying an individual at risk of allergic reaction to

CC an antigen by identifying prior display of allergic symptoms when exposed to the antigen, or a familial relationship with an individual who previously displayed allergic symptoms when exposed to the antigen. Following this an antigen-specific IgE present on one or more mast cells or basophils in the individual's serum is identified. The individual is then contacted with a peptide corresponding to a portion of the antigen, which is selected, formulated, and delivered so that binding of the peptide to antigen-specific IgE is reduced as compared with IgE binding of intact antigen. The composition is also useful for treating and preventing allergic reactions.

XX
SQ Sequence 2041 BP; 637 A; 473 C; 530 G; 401 T; 0 other;

Query Match 100.0%; Score 2032; DB 22; Length 2041;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 aataatcatatatattcatcaatcatctatatataagtagtagcaggagcaatgagggag 60
|||||
Db 1 aataatcatatatattcatcaatcatctatatataagtagtagcaggagcaatgagggag 60
|||||
QY 61 ggtttctccactgatgctgtgtagggatccttgctcctggcttcagtttctgcaacgca 120
|||||
Db 61 ggtttctccactgatgctgtgtagggatccttgctcctggcttcagtttctgcaacgca 120
|||||
QY 121 tgccaagtcatcaccttacccagaagaagaaacagagacccctgcgccagaggtgcctcca 180
|||||
Db 121 tgccaagtcatcaccttacccagaagaagaaacagagacccctgcgccagaggtgcctcca 180
|||||
QY 181 gagttgtcaacaggaacccgatgacttgaagcaaaagggcatgcgagtctcgtgcaccaa 240
|||||
Db 181 gagttgtcaacaggaacccgatgacttgaagcaaaagggcatgcgagtctcgtgcaccaa 240
|||||
QY 241 gctcgagtatgatcctcgttgtgtctatgatcctcagggacacactggcaccaccaacca 300
|||||
Db 241 gctcgagtatgatcctcgttgtgtctatgatcctcagggacacactggcaccaccaacca 300
|||||
QY 301 acgttccctccaggggagcgacacgtggccgcccaacccggagactacgatgatgaccg 360
|||||
Db 301 acgttccctccaggggagcgacacgtggccgcccaacccggagactacgatgatgaccg 360
|||||
QY 361 ccgtcaacccccgaagagaggagggccgatggggaccacactggaccaggagcggtga 420
|||||
Db 361 ccgtcaacccccgaagagaggagggccgatggggaccacactggaccaggagcggtga 420
|||||
QY 421 aagagaagaagactggagacaaccaagaagaatgtggagggcaccagtgcacagcc 480
|||||
Db 421 aagagaagaagactggagacaaccaagaagaatgtggagggcaccagtgcacagcc 480
|||||
QY 481 acggaataaaggcccggaagagagaaagagagaaagagggggaacaccaggtagcca 540
|||||
Db 481 acggaataaaggcccggaagagagaaagagagaaagagggggaacaccaggtagcca 540
|||||
QY 541 tgtgagggaagaaacatctcgggaacaaccccttctacttccgtcaaggcggttagcac 600
|||||
Db 541 tgtgagggaagaaacatctcgggaacaaccccttctacttccgtcaaggcggttagcac 600
|||||
QY 601 ccgtacgggaaccaaaccggttaggatccgggtcctcgtcagaggtttgaccaaaaggtcaag 660
|||||
Db 601 ccgtacgggaaccaaaccggttaggatccgggtcctcgtcagaggtttgaccaaaaggtcaag 660
|||||
QY 661 gcagtttcagaatctccagaatcacccgtatttgtcagatcgaggccaaacacactct 720
|||||
Db 661 gcagtttcagaatctccagaatcacccgtatttgtcagatcgaggccaaacacactct 720
|||||
QY 721 tgttctcccaagcacgctgatgtgataacatccttgttatccagcaaggccac 780
|||||
Db 721 tgttctcccaagcacgctgatgtgataacatccttgttatccagcaaggccac 780
|||||
QY 781 cgtgaccgtagcaaatggcaataacagaaagagctttaatcttgacgagggccatgact 840
|||||
Db 781 cgtgaccgtagcaaatggcaataacagaaagagctttaatcttgacgagggccatgact 840
|||||

QY 841 cagaatcccatcgggtttcatttcctacatcttgaaaccgcatgacaaccagaaacctcag 900
Db 841 cagaatcccatcgggtttcatttcctacatcttgaaaccgcatgacaaccagaaacctcag 900
QY 901 agtagctaaatcccatgcccgttaacacacccggccagtttgaggatttcttcccgcg 960
Db 901 agtagctaaatcccatgcccgttaacacacccggccagtttgaggatttcttcccgcg 960
QY 961 gacgagccgagaccaatcatcctactctcagggccttcagcaggagaatcagttggaggccgc 1020
Db 961 gacgagccgagaccaatcatcctactctcagggccttcagcaggagaatcagttggaggccgc 1020
QY 1021 cttcaatgcggaattcaatgagatacagcagggtgctgttagaagagaatgcaggaggtga 1080
Db 1021 cttcaatgcggaattcaatgagatacagcagggtgctgttagaagagaatgcaggaggtga 1080
QY 1081 gcaagagagagagggcagagggcagtgactcggagtactcggagttagaacaatgaaggagt 1140
Db 1081 gcaagagagagagggcagagggcagtgactcggagtactcggagttagaacaatgaaggagt 1140
QY 1141 gatagctaaagtccaagggagcagctttaagaacttactaagcagcgtataatccgtctc 1200
Db 1141 gatagctaaagtccaagggagcagctttaagaacttactaagcagcgtataatccgtctc 1200
QY 1201 aaagaaaggctccgaagaagaggagatatcaccaccccaatcaacttgagagaagcga 1260
Db 1201 aaagaaaggctccgaagaagaggagatatcaccaccccaatcaacttgagagaagcga 1260
QY 1261 gcccgatcttttaacaaactttgggaagtatttgagggtgaagccagacaagaagacc 1320
Db 1261 gcccgatcttttaacaaactttgggaagtatttgagggtgaagccagacaagaagacc 1320
QY 1321 ccagcttcaggacctggacatgatgctcacctgtgtagagatcaagaaggagctttgat 1380
Db 1321 ccagcttcaggacctggacatgatgctcacctgtgtagagatcaagaaggagctttgat 1380
QY 1381 gctcccacacttcaactcaaggccatggttatcgtcgtcgtcaacaaaggaaactgaaa 1440
Db 1381 gctcccacacttcaactcaaggccatggttatcgtcgtcgtcaacaaaggaaactgaaa 1440
QY 1441 ccttgaactcgtggttaagaaaagagcaacaacagaggggacggcggaagaagagga 1500
Db 1441 ccttgaactcgtggttaagaaaagagcaacaacagaggggacggcggaagaagagga 1500
QY 1501 ggacgaagacgaagaagaggagggaagtaacagagaggtgcgttagtacacagcgaggtt 1560
Db 1501 ggacgaagacgaagaagaggagggaagtaacagagaggtgcgttagtacacagcgaggtt 1560
QY 1561 gaaggaggcgatgtgttcatcatgccagcagctcatccagtagccatcaacgcttcctc 1620
Db 1561 gaaggaggcgatgtgttcatcatgccagcagctcatccagtagccatcaacgcttcctc 1620
QY 1621 cgaactccatctgcttggttcggtatcaacgctgaaacaaacacacagaatcttccttg 1680
Db 1621 cgaactccatctgcttggttcggtatcaacgctgaaacaaacacacagaatcttccttg 1680
QY 1681 aggtgataaggacaatgtgatagaccagatagagaagcaagcgaaagatttagcattccc 1740
Db 1681 aggtgataaggacaatgtgatagaccagatagagaagcaagcgaaagatttagcattccc 1740
QY 1741 tgggtcgggtgaacaagttgagaagctcatcaaaaaaccagaaggaatctcactttgtgag 1800
Db 1741 tgggtcgggtgaacaagttgagaagctcatcaaaaaaccagaaggaatctcactttgtgag 1800
QY 1801 tgctcgtcctcaatctcaatctcaatctcgtcgtctcctgagaaagagtgctcctgagaa 1860
Db 1801 tgctcgtcctcaatctcaatctcaatctcgtcgtctcctgagaaagagtgctcctgagaa 1860
QY 1861 agaggatcaagaggaggaacaaacaaaggagggtccactccttcaattttgaaggg 1920
Db 1861 agaggatcaagaggaggaacaaacaaaggagggtccactccttcaattttgaaggg 1920

QY 1921 ttttaactgagaatgaggcaactgttatgtatcgataataagatcacgcttttgtact 1980
Db 1921 ttttaactgagaatgaggcaactgttatgtatcgataataagatcacgcttttgtact 1980
QY 1981 ctactatccaaaaaacttatcaataaaaaaacggttgcgtgtgtttctcc 2032
Db 1981 ctactatccaaaaaacttatcaataaaaaaacggttgcgtgtgtttctcc 2032

RESULT 3

AAT76613
ID AAT76613 standard; cDNA to mRNA; 2032 BP.

XX AAT76613;

DT 29-DEC-1997 (first entry)

XX Peanut allergen Ara hi cDNA clone P41b.

XX Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy;
KW monoclonal antibody; ELISA; analysis; Ara hi; ds.

OS Arachis hypogaea strain Florunner.

XX Key Location/Qualifiers
FH CDS 50..1930
FT /*tag= a
FT sig_peptide 50..115
FT /*tag= b
FT mat_peptide 116..1927
FT /*tag= c
FT polyA_signal 2005..2010
FT /*tag= d

XX WO9724139-A1.

XX 10-JUL-1997.

XX 23-SEP-1996; 96WO-US15222.

XX 04-MAR-1996; 96US-0610424.

XX 29-DEC-1995; 95US-0009455.

XX (UYAR-) UNIV ARKANSAS.

XX Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;

XX WPI; 1997-363453/33.

XX P-PSDB; AAW22150.

XX Peanut allergens Ara hi and Ara hII - used for vaccination and in
XX two-site monoclonal antibody based ELISA
XX Claim 31; Page 186-188; 354pp; English.

XX This cDNA clone, designated P41b, codes for the major peanut
XX allergen Ara hi (AAW22150), which has multiple IGE binding epitopes
XX (see AAW24165-87). It was amplified from peanut seed cDNA
XX using a primer (see AAT76616) based on an isolated Ara hi peptide
XX (see AAW24206). The sequence shows significant homology with the
XX vicilin family of seed storage proteins of other legumes. The
XX gene is capable of producing a protein product in prokaryotic
XX cells that is recognised by serum IGE from a large proportion of
XX individuals with peanut hypersensitivity. Ara hi and Ara hII (see
XX AAW24164) can be used to raise monoclonal antibodies which are used
XX in a specific two-site MAb ELISA for the detection of Ara hi or Ara
XX hII (claimed). IGE-binding Ara antigen epitopes may be used in
XX vaccines to protect against allergic reactions to peanut allergens,
XX e.g. anaphylactic shock.

XX Sequence 2032 BP; 628 A; 473 C; 530 G; 401 T; 0 other;

Art

| | | | | | | | | | |
|--|-----|--|------|--|--|--|--|--|--|
| Query Match 99.7%; Score 2025.6; DB 18; Length 2032; Best Local Similarity 99.8%; Pred. No. 0; Matches 2028; Conservative 0; Mismatches 4; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 1 | aataatcatatatattcatcaatcatctatatataagtagtagcaggagcaaatgagaggag | 60 | | | | | | |
| Db | 1 | aataatcatatatattcatcaatcatctatatataagtagtagcaggagcaaatgagaggag | 60 | | | | | | |
| QY | 61 | ggtttctccactgatgctgttgctagggatccttgctcctgggttcagtttctgcaacgca | 120 | | | | | | |
| Db | 61 | ggtttctccactgatgctgttgctagggatccttgctcctgggttcagtttctgcaacgca | 120 | | | | | | |
| QY | 121 | tgccaaagtcatcacottaccagaagaataaacagagaacccctgcgccagaggtgcctcca | 180 | | | | | | |
| Db | 121 | tgccaaagtcatcacottaccagaagaataaacagagaacccctgcgccagaggtgcctcca | 180 | | | | | | |
| QY | 181 | gagttgtcaacaggaaccgggatgacttgaagcaaaaaggcatgcgactcctgcgtgcacca | 240 | | | | | | |
| Db | 181 | gagttgtcaacaggaaccgggatgacttgaagcaaaaaggcatgcgactcctgcgtgcacca | 240 | | | | | | |
| QY | 241 | gctcgagtatgatcctcgttgtgtctatgatcctcgaggacacactggcaccaccaacca | 300 | | | | | | |
| Db | 241 | gctcgagtatgatcctcgttgtgtctatgatcctcgaggacacactggcaccaccaacca | 300 | | | | | | |
| QY | 301 | acgttccctccaggggagcggacacgtggccgccaaacccggagactacgatgatgaccg | 360 | | | | | | |
| Db | 301 | acgttccctccaggggagcggacacgtggccgccaaacccggagactacgatgatgaccg | 360 | | | | | | |
| QY | 361 | ccgtcaaccccggaagaggaaggagccgatgggaccagctggaccgaggagcgtga | 420 | | | | | | |
| Db | 361 | ccgtcaaccccggaagaggaaggagccgatgggaccagctggaccgaggagcgtga | 420 | | | | | | |
| QY | 421 | aagagaagaagactggagacaaaccaaagagaagattggaggcgaccaaagtcatcagcagcc | 480 | | | | | | |
| Db | 421 | aagagaagaagactggagacaaaccaaagagaagattggaggcgaccaaagtcatcagcagcc | 480 | | | | | | |
| QY | 481 | acggaaaaataaggcccgaaaggaagagaaggagaaacaagagtggggaacaccaggtagcca | 540 | | | | | | |
| Db | 481 | acggaaaaataaggcccgaaaggaagagaaggagaaacaagagtggggaacaccaggtagcca | 540 | | | | | | |
| QY | 541 | tgtgagggaagaaacatctcggaaacaacccctttctacttcccgtaaggcggttttagcac | 600 | | | | | | |
| Db | 541 | tgtgagggaagaaacatctcggaaacaacccctttctacttcccgtaaggcggttttagcac | 600 | | | | | | |
| QY | 601 | ccgctacggggaacccaaaacggtaggatccgggtcctgcagaggtttgaccaaaggtcaag | 660 | | | | | | |
| Db | 601 | ccgctacggggaacccaaaacggtaggatccgggtcctgcagaggtttgaccaaaggtcaag | 660 | | | | | | |
| QY | 661 | gcagtttcagaatctccagaatcacccgtattgtgcagatcgaggccaaacctaacactct | 720 | | | | | | |
| Db | 661 | gcagtttcagaatctccagaatcacccgtattgtgcagatcgaggccaaacctaacactct | 720 | | | | | | |
| QY | 721 | tgttcttcccaagcaacgctgatgctgataacatccttgtttatccagcaagggcaagccac | 780 | | | | | | |
| Db | 721 | tgttcttcccaagcaacgctgatgctgataacatccttgtttatccagcaagggcaagccac | 780 | | | | | | |
| QY | 781 | cgtgaccgtagcaaaatggcaataaacagaaagagctttaatcttgacgagggccatgcact | 840 | | | | | | |
| Db | 781 | cgtgaccgtagcaaaatggcaataaacagaaagagctttaatcttgacgagggccatgcact | 840 | | | | | | |
| QY | 841 | cagaatcccatcccggtttcatttccatatttgaaacccgccaatgcacaaccagaaacctcag | 900 | | | | | | |
| Db | 841 | cagaatcccatcccggtttcatttccatatttgaaacccgccaatgcacaaccagaaacctcag | 900 | | | | | | |
| QY | 901 | agtagctaaaaatctccatgcccggttaacacacccggccagtttgaggatttcttcccggc | 960 | | | | | | |
| Db | 901 | agtagctaaaaatctccatgcccggttaacacacccggccagtttgaggatttcttcccggc | 960 | | | | | | |
| QY | 961 | gagcagccgagaccaatcatcctacttgacgggcttcagcagggaatacgttggaggccgc | 1020 | | | | | | |
| Db | 961 | gagcagccgagaccaatcatcctacttgacgggcttcagcagggaatacgttggaggccgc | 1020 | | | | | | |

| | | | |
|----|------|--|------|
| QY | 1021 | cttcaatcggaattcaatgagatacggagggtgctgttagaagagaaatgcaggagggtga | 1080 |
| Db | 1021 | cttcaatcggaattcaatgagatacggagggtgctgttagaagagaaatgcaggagggtga | 1080 |
| QY | 1081 | gcaagaggagagagggcgagggcgatggagtagtctcggagtagtgagaaacaatgaaggagt | 1140 |
| Db | 1081 | gcaagaggagagagggcgagggcgatggagtagtctcggagtagtgagaaacaatgaaggagt | 1140 |
| QY | 1141 | gatagtcaaaagtgtcaaaaggagcacggttgaagaacttactaaagcacgcgtataatccgtctc | 1200 |
| Db | 1141 | gatagtcaaaagtgtcaaaaggagcacggttgaagaacttactaaagcacgcgtataatccgtctc | 1200 |
| QY | 1201 | aaagaaaggctccgaagaagaaggagagatatcaccaacccaatcaacttgagagaaaggcga | 1260 |
| Db | 1201 | aaagaaaggctccgaagaagaaggagagatatcaccaacccaatcaacttgagagaaaggcga | 1260 |
| QY | 1261 | gccgatcttttctaacaactttgggaagtattattgagggtgaagccagacaagaagaaccc | 1320 |
| Db | 1261 | gccgatcttttctaacaactttgggaagtattattgagggtgaagccagacaagaagaaccc | 1320 |
| QY | 1321 | ccagcttcaggacctggacatgatgctcacctgtgtagagatcaaaagaaggagcgtttgat | 1380 |
| Db | 1321 | ccagcttcaggacctggacatgatgctcacctgtgtagagatcaaaagaaggagcgtttgat | 1380 |
| QY | 1381 | gctccacacacttcaactcaaaaggccatggttatactcgtcgtcaacaaagggaactggaaa | 1440 |
| Db | 1381 | gctccacacacttcaactcaaaaggccatggttatactcgtcgtcaacaaagggaactggaaa | 1440 |
| QY | 1441 | ccttgaactcgtggctgtaagaaaaagacaacagaggggacgcgcggaagaagagga | 1500 |
| Db | 1441 | ccttgaactcgtggctgtaagaaaaagacaacagaggggacgcgcggaagaagagga | 1500 |
| QY | 1501 | ggacgaagacgaagaagagggaagtaaacagagaggtgcgtaggtacacagcgaggtt | 1560 |
| Db | 1501 | ggacgaagacgaagaagagggaagtaaacagagaggtgcgtaggtacacagcgaggtt | 1560 |
| QY | 1561 | gaaggaaaggcgatgtgttcacatcatgccagcagctccagtagccatcaacgcttcctc | 1620 |
| Db | 1561 | gaaggaaaggcgatgtgttcacatcatgccagcagctccagtagccatcaacgcttcctc | 1620 |
| QY | 1621 | cgaactccatctcgttggcttggtatcaacgctgaacaaacacacagaaatcttctctgc | 1680 |
| Db | 1621 | cgaactccatctcgttggcttggtatcaacgctgaacaaacacacagaaatcttctctgc | 1680 |
| QY | 1681 | aggtgataaggacaatgtgatagaccagatagagaagcaagcgaaagatttagcattccc | 1740 |
| Db | 1681 | aggtgataaggacaatgtgatagaccagatagagaagcaagcgaaagatttagcattccc | 1740 |
| QY | 1741 | tgggtcggtgaaacaaagttgaaagctcatcaaaaacagaaaggaatctcactttgtgag | 1800 |
| Db | 1741 | tgggtcggtgaaacaaagttgaaagctcatcaaaaacagaaaggaatctcactttgtgag | 1800 |
| QY | 1801 | tgctcgtcctcaatctcaatctcctcgtcgtctcctgagaaagagtagtctcctgagaa | 1860 |
| Db | 1801 | tgctcgtcctcaatctcaatctcctcgtcgtctcctgagaaagagtagtctcctgagaa | 1860 |
| QY | 1861 | agaggatcaagaggaggaaccaaaggagggaaggtccactcctttccaatttgaaggc | 1920 |
| Db | 1861 | agaggatcaagaggaggaaccaaaggagggaaggtccactcctttccaatttgaaggc | 1920 |
| QY | 1921 | ttttaactgagaatggaggcaacttgttatgtatcgataataagatcacgcttttgtact | 1980 |
| Db | 1921 | ttttaactgagaatggaggcaacttgttatgtatcgataataagatcacgcttttgtact | 1980 |
| QY | 1981 | ctactatccaaaaaacttatcaataataaaaaaacggttgtgctgtttctcc 2032 | |
| Db | 1981 | ctactatccaaaaaacttatcaataataaaaaaacggttgtgctgtttctcc 2032 | |

CC antigenic peptides having a reduced ability to bind immunoglobulin E
CC (IgE) as compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the
CC antigen, which is selected, formulated, and delivered so that binding of
CC the peptide to antigen-specific IgE is reduced as compared with IgE
CC binding of intact antigen. The composition is also useful for treating
CC and preventing allergic reactions.

XX
SQ Sequence 1952 BP; 608 A; 456 C; 512 G; 376 T; 0 other;

Query Match 86.1%; Score 1749.8; DB 22; Length 1952;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 42; Indels 58; Gaps 6;

| | | | |
|----|-----|--|-----|
| QY | 48 | caatgagaggagggttctccactgatgctgttgctagggatccttgtcctggttcag | 107 |
| Db | 1 | caatgagaggagggttctccactgatgctgttgctagggatccttgtcctggttcag | 60 |
| QY | 108 | ttctgcaacgcatgccaaagtcatcaccttaccagaagaaaacagagaaacccctgcgcc | 167 |
| Db | 61 | ttctgcaacgcatgccaaagtcatc-----ttaccggaaaacagagaaacccctgcgcc | 114 |
| QY | 168 | agaggtgcctccagagttgtcaacaggaaacccggatgacttgaagcaaaaggcatgcgagt | 227 |
| Db | 115 | agaggtgcctccagagttgtcaacaggaaacccggatgacttgaagcaaaaggcatgcgagt | 174 |
| QY | 228 | ctcgctgcaccaagctcgagtatgatccctggtgtgtctatgatcctcgaggacacactg | 287 |
| Db | 175 | ctcgctgcaccaagctcgagtatgatccctggtgtgtctatg-----cacactg | 223 |
| QY | 288 | gcaccaccaaccaacgttccctccaggggagcgagacacgtggccgcgaaccccgagact | 347 |
| Db | 224 | gcgcaccaaccaacgtcacccctccaggggagcgagacacgtggccgcgaaccccgagact | 283 |
| QY | 348 | acgatgatgacccgcctcaaccccgaaagagaggaagccgatggggaccagctggac | 407 |
| Db | 284 | acgatgatgacccgcctcaaccccgaaagagaggaagccgatggggaccagctggac | 343 |
| QY | 408 | cgagggagcgtgaaagagaagaagactggagacaaaccaagaagaagattggagcgacca | 467 |
| Db | 344 | cgagggagcgtgaaagagaagaagactggagacaaaccaagaagaagattggagcgacca | 403 |
| QY | 468 | gtcatcagcagccacggaaaaataaggcccgaaagagaagaaggaacaagagtgggaa | 527 |
| Db | 404 | gtcatcagcagccacggaaaaataaggcccgaaagagaagaaggaacaagagtgggaa | 463 |
| QY | 528 | caccaggtagccatgtgagggagaagaacatctcggaacaaccccttctacttcccgtaa | 587 |
| Db | 464 | caccaggtagcaggtgagggagaagaacatcacggaaacaccccttctacttcccgtaa | 523 |
| QY | 588 | ggcggttagcaccgcctacgggaaccaaagcgttaggatccgggtcctgcagaggttg | 647 |
| Db | 524 | ggcggttagcaccgcctacgggaaccaaagcgttaggatccgcgcctgcagaggttg | 583 |
| QY | 648 | accaaaggtcaaggcagtttcagaatctccagaatcacccgtattgtgcagatcgagcca | 707 |
| Db | 584 | accaaaggtcaaggcagtttcagaatctccagaatcacccgtattgtgcagatcgagcca | 643 |
| QY | 708 | aacctaacactcttgttcttcccaagcacgctgatgctgataaacatcctgttatccagc | 767 |
| Db | 644 | gacctaacactcttgttcttcccaagcacgctgatgctgataaacatcctgttatccagc | 703 |

| | | | |
|----|------|---|------|
| QY | 768 | aagggaagccaccgtgaccgtagcaaatggcaataacagaaagagctttaatcttgacg | 827 |
| Db | 704 | aaggacaagccaccgtgaccgtagcaaatggcaataacagaaagagctttaatcttgacg | 763 |
| QY | 828 | agggccatcactcagaatcccatccggtttcatttctacatcttgaaccccatgaca | 887 |
| Db | 764 | agggccatcactcagaatcccatccggtttcatttctacatcttgaatcgacatgaca | 823 |
| QY | 888 | accagaacctcagatgactaaatctccatgccccgttaacacaccccgccagtttgag | 947 |
| Db | 824 | accagaacctcagatgactaaatctccatgccccgttaacacaccccgccagtttgag | 883 |
| QY | 948 | atttctcccgcgagcagccgagaccaatcatcctacttgcagggcttcagcaggaata | 1007 |
| Db | 884 | atttctcccgcgagcagccgagaccaatcatcctacttgcagggattcagcaggaata | 943 |
| QY | 1008 | cgttgaggccgccttcaatgcggaattcaatgagatacggagggtgctgttagaagaga | 1067 |
| Db | 944 | ctttggaggccgccttcaatgcggaattcaatgagatacggagggtgctgttagaagaga | 1003 |
| QY | 1068 | atgcaggagggtgagcaagagagagagggcgagagggcgatggagttactcggagttagt | 1127 |
| Db | 1004 | atgcaggaggagagcaagagagagagagggcgagagggcgagttactcggagttagtg | 1060 |
| QY | 1128 | acaatgaaggagtgatagtcaaaaggtgcaaaaggagcaggttgaaagaacttactaagc | 1187 |
| Db | 1061 | ataatgaaggagtgatagtcaaaaggtgcaaaaggagcaggttcaagaacttactaagc | 1120 |
| QY | 1188 | ctaaatccgtctcaaaagaaaggctccgaagaagagggagatatcaccaacccaatca | 1247 |
| Db | 1121 | ctaaatccgtctcaaaagaaaggctccgaagaaga---agatatcaccaacccaatca | 1177 |
| QY | 1248 | tgagagaaggcgagcccgatcttttaacaaactttgggaagttatttgaggtgaagccag | 1307 |
| Db | 1178 | tgagagaaggcgagcccgatcttttaacaaactttgggaagttatttgaggtgaagccag | 1237 |
| QY | 1308 | acaagaagaaccccccagcttcaggacctggacatgatgtcacctgtgtagatgataag | 1367 |
| Db | 1238 | acaagaagaaccccccagcttcaggacctggacatgatgtcacctgtgtagatgataag | 1297 |
| QY | 1368 | aaggagctttgatgtcccccacacttcaactcaaaagggccatggttatcgtcgtgtaaca | 1427 |
| Db | 1298 | aaggagctttgatgtcccccacacttcaactcaaaagggccatggttatcgtcgtgtaaca | 1357 |
| QY | 1428 | aaggaactgaaacacttgaaactcgtggctgtgaagaaaagagcaacaacagagggcg | 1487 |
| Db | 1358 | aaggaactgaaacacttgaaactcgtagctgtgaagaaaagagcaacaacagagggcg | 1417 |
| QY | 1488 | gggaa-----gaagaggaggaggaagacgaagaagagggaggaagtaacagagaggt | 1539 |
| Db | 1418 | gggaacaagaggtcgaaagagaaggaggaagactaagaagagggaggaagtaacagagaggt | 1477 |
| QY | 1540 | gcgtaggtacacagcgaggttgaaagagggcgatgtgttcatcatgccagcagctcatcc | 1599 |
| Db | 1478 | gcgtaggtacacagcgaggttgaaagagggcgatgtgttcatcatgccagcagctcatcc | 1537 |
| QY | 1600 | agtagccatcaacgcttccctccgaactccatctgcttggcttcggtatcaacgctgaaaa | 1659 |
| Db | 1538 | agtagccatcaacgcttccctccgaactccatctgcttggcttcggtatcaacgctgaaaa | 1597 |
| QY | 1660 | caaccacagaatcttcttgcaggtgataaggacaatgtgatagaccagatagagaagca | 1719 |
| Db | 1598 | caaccacagaatcttcttgcaggtgataaggacaatgtgatagaccagatagagaagca | 1657 |
| QY | 1720 | agcgaaggatttagcattccctgggtcggttgaaacaagtttgagaagctcatcaaaaaacca | 1779 |
| Db | 1658 | agcgaaggatttagcattccctgggtcggttgaaacaagtttgagaagctcatcaaaaaacca | 1717 |
| QY | 1780 | gaaggaaatctcaactttgtgagtgctcgtcctcaatctcaatctcaatctcgtctctcc | 1839 |
| Db | 1718 | gagggagctcaactttgtgagtgctcgtcctcaatctcgtcctcaatctcgtc----- | 1759 |
| QY | 1840 | tgagaaaagagttctcctgagaaaaggaggtatcaagagggaggaacccaagggaggtcc | 1899 |

Db 1760 -----gtctcctgaaaagaggatcaagaggaggaaaacccaaggagggaagggtcc 1810
QY 1900 actccttcaatttgaaggcttttaactgagaatggaggcaacttggtatgtatcgata 1959
Db 1811 actccttcaatttgaaggcttttaactgagaatggaggaaacttggtatgtatccata 1870
QY 1960 ataagatcacgctttgtactactactatccaaaaacttatcaataataaaaaacgtttgt 2019
Db 1871 ataagatcacgctttgtaatctactatccaaaaacttatcaataataaaaaacgtttgt 1930
QY 2020 gcgttggtttctcc 2032
Db 1931 gcgttggtttctcc 1943

RESULT 7
AAV17564
AAV17564 standard; cDNA; 1320 BP.
XX AAV17564;
AC
XX
DT 10-JUN-1998 (first entry)
XX
DE Coding sequence for the beta subunit of beta-conglycinin.
XX Beta-conglycinin; soybean seed protein; transgenic plant;
KW seed storage protein profile; ss.
KW
XX Glycine max.
OS
XX
PN W09747731-A2.
XX
PD 18-DEC-1997.
XX
PF 10-JUN-1997; 97WO-US09743.
XX
PR 14-JUN-1996; 96US-0019940.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Fader GM, Kinney AJ;
XX
DR WPI; 1998-052298/05.
XX
PT Suppression of specific classes of soybean seed protein genes -
PT useful to change seed storage protein profiles of transgenic plants
XX
PS Disclosure; Page 31-32; 58pp; English.
XX
CC This sequence represents the coding sequence for the beta subunit of the
CC soybean seed protein beta-conglycinin. The method of the invention is for
CC reducing the quantity of a soybean seed storage protein (A), such as
CC beta-conglycinin, in soybeans. It comprises: (a) constructing a chimeric
CC gene comprising: (i) a nucleic acid fragment encoding a promoter that is
CC functional in the cells of soybean seeds; (ii) a nucleic acid fragment
CC encoding all or a portion of (A) placed in sense or antisense orientation
CC relative to the promoter of (i); and (iii) a transcriptional termination
CC region; (b) creating a transgenic soybean cell by introducing into a
CC soybean cell the chimeric gene of (a); and (c) growing the transgenic
CC soybean cells of (b) under conditions that result in expression of the
CC chimeric gene of (a); where the quantity of one or more members of a
CC class of (A) subunits is reduced when compared to soybeans not containing
CC the chimeric gene of (a). The method is used to construct transgenic
CC soybean lines where the expression of genes encoding (A) are modulated to
CC effect a change in seed storage protein profile of transgenic plants.
CC Modification of the seed storage protein profile can result in the
CC production of novel soy protein products with unique and valuable
CC functional characteristics.
XX
SQ Sequence 1320 BP; 394 A; 315 C; 285 G; 326 T; 0 other;

Query Match 20.0%; Score 406.8; DB 19; Length 1320;
Best Local Similarity 62.1%; Pred. No. 2.7e-110;
Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;

QY 562 gaacaaacccctttctactt---cccgtcaagcggttttagcaccgcgtacgggaacccaaaa 618
Db 90 gaataacccctttctactttagaagctctaaacgcttccaaactctctttgagaaacccaaa 149

QY 619 cgttaggatccgggtcctgcagaggttttgacaaaagtgcaaggcaggtttcagaatatccca 678
Db 150 cgttcgcattcgtctcctccagagattcaaaaacgctcccaaacacttgagaaccttcg 209

QY 679 gaatcaccggtattgtgcagatcgaggcccaaacctaaacactcttcttcccaagcagc 738
Db 210 agactaccggattgtccagtttcagtcaaaaccccaacacaaatctcttccccaccatgc 269

QY 739 tgatgctgataacatccctgtttatccagcaaggggccatgcactcagaatccccatcggtt 798
Db 270 tgacgcgatttctcctcctttgtccttagcgggagagccatacttaccttggtgaacaa 329

QY 799 caataacagaaagagctttaactcttgacgagggccatgcactcagaatccccatcggttt 858
Db 330 cgacgacagagactcctacaaccttcacctggcgatgccagagaatccccagctggaac 389

QY 859 catttctacatcttgaaccgccatgacaaacagaaacctcagagtagctaaaaatctccat 918
Db 390 cacttactatttggttaacccttcacgaccaccagaatctcaaaaataatcaaaacttgccat 449

QY 919 gcccgtaaacaaccccgccagtttgaggatttcttccccgcgagcagccgagaccaatc 978
Db 450 acccgtaacaacacctggcagatatgatgatttcttcttatctagcactcaagcccaaca 509

QY 979 atctacttgcagggcttcagcaggaatacgttggaggccgcttcaatcggaattcaa 1038
Db 510 gtctacttgcaggcttcagccataatattcttagagacctcttccatagcgaattcga 569

QY 1039 tgagatacggagggtgctgttagaagagaatgcaggaggtgagcaagaggagagagggca 1098
Db 570 ggagataaacagggttttgttggagaggaaggagcagagggcagcaagag----- 621

QY 1099 gaggcgatggagtactcggagtagtgagaaacaatgaaggagtgatagtcaaaagtgtcaaa 1158
Db 622 -----ggagtgcgtggaactctcaaa 644

QY 1159 ggagcacgttgaaagaacttactaagcacgtctaaatccgtctcaaaagaaaggctccgaaga 1218
Db 645 ggaacaaattcggcaactgcagcagacgtgcaaatctagtccaaggaaacccattctctc 704

QY 1219 agagggagatatcaccaacccaatcaacttgagagaaggcgagcccgatcttcttaacaa 1278
Db 705 cgaagatg-----aaccattcaacttgagaagcgcgaaccccatctattccaacaa 755

QY 1279 ctttgggaagtatttgagggtgaagccagacaagaagaaccccgacttcaggagcctgga 1338
Db 756 ctttggaaaagtcttcttgagatcacccctg--agaaaaacccacagcttcgggacttggga 812

QY 1339 catgatgctcacctgtgtagagatcaaaagaaggagcttttgatctcccacacttcaactc 1398
Db 813 tatcttctcagttctgtggatatcaacgaaggagctcttcttaccacacttcaattc 872

QY 1399 aaagcccatggttatcgtcgtcgtcaacaaaggaaactggaaaaccttgaaactcgtggctgt 1458
Db 873 aaagcccatagtgtatactagtattaatgaaggagatgcaaacattgaaacttggtggcat 932

QY 1459 aagaaaaagacaacaacagaggggacggcggaagaagaaggaggacgaagacgaagaaga 1518
Db 933 taaagaacaacaacagagagaaacagggaagaggaaaccttg----- 975

QY 1519 ggagggaagtaacagagaggtgcgtaggtacacagcgaggttgaaggagcgatgtgtt 1578
Db 976 -----gaagtgcgaagggtacagagctgaattgtctgaagacgatgtatt 1019

QY 1579 catcatgccagcagctcatccagtagccatcaacgcttctcctcgaactccatctctgttg 1638

Db 1020 tgtaattccagcagcgtattatccatttgtgtgtaacgctacctcaaacctcaatttcttgc 1079
QY 1639 cttcgggtatcaacgctgaaaacaaccacagaatcttcccttgaggtgataagacaatgt 1698
Db 1080 ttttgtatcaatgctgagaacaaccagaggaacttcccttgaggcgagaagacaatgt 1139
QY 1699 gatagaccagatagagaagcaagcgaaggatttagcattccctgggtcggtgaacaagt 1758
Db 1140 ggttaaggcagatagaagacaagtgcaggagcttgcgttccctgggtcgcacaagatgt 1199
QY 1759 tgagaagctcatcaaaaaaccagaagggaatctcactttgtgagtgctcgtcctca 1812
Db 1200 tgagaggctattaagaagcagagggaatcctacttctgtgtgatgctcagcctca 1253

RESULT 8
AAV17562
ID AAV17562 standard; cDNA; 1818 BP.
XX
AC AAV17562;
XX
DT 10-JUN-1998 (first entry)
XX
DE Coding sequence for the alpha subunit of beta-conglycinin.
XX
KW Beta-conglycinin; soybean seed protein; transgenic plant;
KW seed storage protein profile; ss.
XX
OS Glycine max.
XX
PN WO9747731-A2.
XX
PD 18-DEC-1997.
XX
PF 10-JUN-1997; 97WO-US09743.
XX
PR 14-JUN-1996; 96US-0019940.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Fader GM, Kinney AJ;
XX
DR WPI; 1998-052298/05.
XX
PT Suppression of specific classes of soybean seed protein genes -
PT useful to change seed storage protein profiles of transgenic plants
XX
PS Disclosure; Page 29-30; 58pp; English.
XX
CC This sequence represents the coding sequence for the alpha subunit of the
CC soybean seed protein beta-conglycinin. The method of the invention is for
CC reducing the quantity of a soybean seed storage protein (A), such as
CC beta-conglycinin, in soybeans. It comprises: (a) constructing a chimeric
CC gene comprising: (i) a nucleic acid fragment encoding a promoter that is
CC functional in the cells of soybean seeds; (ii) a nucleic acid fragment
CC encoding all or a portion of (A) placed in sense or antisense orientation
CC relative to the promoter of (i); and (iii) a transcriptional termination
CC region; (b) creating a transgenic soybean cell by introducing into a
CC soybean cell the chimeric gene of (a); and (c) growing the transgenic
CC soybean cells of (b) under conditions that result in expression of the
CC chimeric gene of (a); where the quantity of one or more members of a
CC class of (A) subunits is reduced when compared to soybeans not containing
CC the chimeric gene of (a). The method is used to construct transgenic
CC soybean lines where the expression of genes encoding (A) are modulated to
CC effect a change in seed storage protein profile of transgenic plants.
CC Modification of the seed storage protein profile can result in the
CC production of novel soy protein products with unique and valuable
CC functional characteristics.
XX
SQ Sequence 1818 BP; 581 A; 424 C; 427 G; 386 T; 0 other;

Query Match 19.9%; Score 404; DB 19; Length 1818;
Best Local Similarity 61.2%; Pred. NO. 2.2e-109;
Matches 765; Conservative 0; Mismatches 410; Indels 75; Gaps 4;

QY 563 aacaaccctttctacttcccgtaaggcgggttagcacccgctacgggaaccaaacggt 622
Db 577 aagaacccttttctcttcggcttaacaggttcgaactcttcaaaaaccaaatacgt 636
QY 623 aggatccgggtcctgcagaggtttgaccaaaggccaaggcagtttcagaatctccagaat 682
Db 637 cgcattcggtcctccagaggttcaaccaacgcctcccaacttcagaatctccgagac 696
QY 683 caccgtattgtgcagatcgaggccaaacctaactcttgttcttcccaagcacgctgat 742
Db 697 taccgcattttggagttcaactccaacccaacacctcttctcccaaccatgctgac 756
QY 743 gctgataaacatccttgttatccagcaaggcggaagccaccgctgacccgtagcaaaatggcaat 802
Db 757 gctgattacctcatcgttatccttaacgggactgcattcttcttcttggtgaacaacgac 816
QY 803 aacagaaaagagctttaatcttgacgagggcccatgcaactcagaatcccccggttttcatt 862
Db 817 gacagagaactcctacagacttcaatctggtgatgcctgagagtcctccctcaggaaaccaca 876
QY 863 tctacatcttgaaccgcccatacgtaacaccagaacctgaagtagtctaaaaatctccatgcc 922
Db 877 tactatgtgtcaaccctgacaacaacgaaaaatctcagattaataaacactcgcataccc 936
QY 923 gttaacacacccggccagtttgaggatttcttcccgcgagcagccgagaccaatcatcc 982
Db 937 gttaacaagcctggtagatttgagagtttcttctctatctagcactgaagctcaacaatcc 996
QY 983 tacttgagggtcttcagcaggaatacgtttggaggccgcttcaatgcgggaattccaatgag 1042
Db 997 tacttgcaaggattcagcaggaacatttttagagggcctctacgataccaacttcgaggag 1056
QY 1043 atacgggagggtgctgttagaagagaatgcaggaggttgagcaagaggagagggcgagagg 1102
Db 1057 ataaacaaggttctgttttagtagaggaaggcgagcagcaaggggagcagaggctg--- 1113
QY 1103 cgatggagtactcggagtgtgagaacaatgaaggagtgatagtcaaaagtgctcaaaaggag 1162
Db 1114 -----caagagagcgtgattgtggaaatctcgaaggaa 1146
QY 1163 cactgtgaagaacttactaaagcacgctaaatccgtctcaaaagaaaggctcogaagaagag 1222
Db 1147 cagattcgggcactgagcaaacgtgccaaatctagttcaaggaaacccatttcttctgaa 1206
QY 1223 ggagatatcaccacaacccaacttagagagaaggcgagcccgatcttcttaacaacttt 1282
Db 1207 g-----ataaaccttttaacttgagaagcgcgaccctactctcaacaacgctt 1257
QY 1283 gggaaagtattattgaggtgaaggcagacagaagaacccccagcttcaggacctggacatg 1342
Db 1258 ggcaagtcttttgagatcaccccgag--agaaaaacccccagcttcgggacttgatatc 1314
QY 1343 atgctcacctgtgtagatcaagaaggagagcttttgatgctccacactcaactcaaaag 1402
Db 1315 ttcctcagtattgtggatatgaacgagggagctcttcttctaccacacttcaattcaag 1374
QY 1403 gccatggttatcgtcgtctcaacaagaagaaactggaaaccttgaactcgtggctgtaaga 1462
Db 1375 gcgatatgatactggttaattaatgaaggagatgcacaacattgaacttgttggccataaa 1434
QY 1463 aaagagcaacaacagaggggacgcggggaagaagaagagagagacgaagcaagaagaggag 1522
Db 1435 gaacaacaacagagcagcaacacaggaagacaccttg----- 1473
QY 1523 ggaagtaacagagaggtgcgtaggttacacagcgaggttgaaaggcgatgtgttcac 1582
Db 1474 -----gaagtgcggaaaatatagagccgaattgtctgacaagaatatatttga 1521
QY 1583 atgccagcagctcatccagtagccatcaacgcgttctcctcgaactccatctctgtgcttc 1642

Db 1522 atcccagcaggttatccagttgtgtcaggtcgtcaacgctacctcaaatctgaatttctgtctatt 1581
QY 1643 ggtatcaacgctgaaaacaaccacagaatcttctctgcaggtgataagacaatgtgata 1702
Db 1582 ggtattaatgccgagaaacaaccagaggaaacttctctgcaggttcgcaagacaatgtgata 1641
QY 1703 gaccagatagagaagcaagcgaaggatttagcattccctgggtcgggtgaaacaagttag 1762
Db 1642 agccagataccttagtcaagtgcaggagcttgctcattccctgggtctgcacaagctgttag 1701
QY 1763 aagctcatcaaaaaaccagaaaggaatctcactttgtgagtgctcgtctca 1812
Db 1702 aagctattaagaagacaaagagaatcctactttgtggatgctcagcctaa 1751

RESULT 9
AAV17563
ID AAV17563 standard; cDNA; 1920 BP.
XX
AC AAV17563;
XX
DT 10-JUN-1998 (first entry)
XX
DE Coding sequence for the alpha' subunit of beta-conglycinin.
XX
KW Beta-conglycinin; soybean seed protein; transgenic plant;
KW seed storage protein profile; ss.
XX
OS Glycine max.
XX
PN W09747731-A2.
XX
PD 18-DEC-1997.
XX
PF 10-JUN-1997; 97WO-US09743.
XX
PR 14-JUN-1996; 96US-0019940.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Fader GM, Kinney AJ;
XX
DR WPI; 1998-052298/05.
XX
PT Suppression of specific classes of soybean seed protein genes -
PT useful to change seed storage protein profiles of transgenic plants
XX
PS Disclosure; Page 30-31; 58pp; English.
XX

CC This sequence represents the coding sequence for the alpha' subunit of
CC the soybean seed protein beta-conglycinin. The method of the invention is
CC for reducing the quantity of a soybean seed storage protein (A), such as
CC beta-conglycinin, in soybeans. It comprises: (a) constructing a chimeric
CC gene comprising: (i) a nucleic acid fragment encoding a promoter that is
CC functional in the cells of soybean seeds; (ii) a nucleic acid fragment
CC encoding all or a portion of (A) placed in sense or antisense orientation
CC relative to the promoter of (i); and (iii) a transcriptional termination
CC region; (b) creating a transgenic soybean cell by introducing into a
CC soybean cell the chimeric gene of (a); and (c) growing the transgenic
CC soybean cells of (b) under conditions that result in expression of the
CC chimeric gene of (a); where the quantity of one or more members of a
CC class of (A) subunits is reduced when compared to soybeans not containing
CC the chimeric gene of (a). The method is used to construct transgenic
CC soybean lines where the expression of genes encoding (A) are modulated to
CC effect a change in seed storage protein profile of transgenic plants.
CC Modification of the seed storage protein profile can result in the
CC production of novel soy protein products with unique and valuable
CC functional characteristics.

XX Sequence 1920 BP; 634 A; 444 C; 449 G; 393 T; 0 other;
SQ

Query Match 16.0%; Score 324.2; DB 19; Length 1920;
Best Local Similarity 58.5%; Pred. No. 1.2e-85;
Matches 764; Conservative 0; Mismatches 408; Indels 135; Gaps 6;
QY 563 aacaacccctttctacttcccgtcaaggcgggttttagcaccccgctacgggaaccaaagcgtt 622
Db 625 aagaacccctttctacttcaactcaaaaggttccaaactcttcaaaaaccaatatggc 684
QY 623 aggatccgggtcctgcagaggtttgacaaaaggtcaaggcaggtttcagaatctccagaat 682
Db 685 cacgttcgctcctccagaggttcaacaaacgctcccaacagcttcagaatctccgagac 744
QY 683 caccgtattgtcagatcgaggccaaacctaactcttcttctcccaagcaecgtgat 742
Db 745 taccgcattttggagttcaactccaacccaacaccccttcttctccccaccatgctgac 804
QY 743 gctgataacatccttgttatccagcaaggcgcaagccacccgtgaccgtagcaaatgggcaat 802
Db 805 gctgattacctcatcgttatccttaacgggagctgccattcttacccttggtagaacaacgac 864
QY 803 aacagaaaagagcctttaatcttgacgaggggccatgcactcagaatccc----- 849
Db 865 gaccgagactcttacaaaccttcaatctggcgatgccctaagagtccttcgcaggaaaccaca 924
QY 850 -----atccgggtttcatttcc 865
Db 925 ttctatgtgtttaaccctgacaacgacgagaaatctcagaatgatagcaggaaaccacattc 984
QY 866 tacatcttgaaccgcccatgacaacacgagaacctcagagtagttagtaaaatctccatgcccgtt 925
Db 985 tatgtgttaaccctgacaacgacgagaatctcagaatgataacactcgccatacccgtt 1044
QY 926 aacacaccccgccagtttgaggatttcttcccgcgagcagccgagacccaatcctctac 985
Db 1045 acaaaacccggtagatttgagagtttcttctctatcttagcaactcaagctcaacagtcctac 1104
QY 986 ttgcagggttcacgaggaatacgttggaggccgcttcaatcggaattcaatgagata 1045
Db 1105 ttgcaagggttcagcaagaatatcttagaggcctcatacagacaccaaatctcgaggagata 1164
QY 1046 cggagggtgctgttagaagagaaatcgaggaggtgagcaagaggagagagggcagaggcga 1105
Db 1165 acaaggttctgttgtagagagggggagcaacaaggaggagaggtgc----- 1219
QY 1106 tggagtactcggagtagtgagaacaatgaaggagtgatagtcaaaagtgtcaaaaggagcac 1165
Db 1220 -----aagagagtgatgtgtgaaatctcaaaagaaacaa 1254
QY 1166 gttgaagaacttactaagcacgcgtaaatccgtctcaaaagaaaggctccgaaagaggga 1225
Db 1255 attcgggaactgagcaaacatgccaaatctagttcaaggaaaaaccatttcttctgaag-- 1312
QY 1226 gatataccaacccaatcaacttgagagaaaggcgagcccgatcttcttaacaactttggg 1285
Db 1313 -----ataaacctttcaacttgggaagccgcgaccccatctattccaacaagcttggc 1365
QY 1286 aagttatttgaggtgaagccagacagaagaacccccagcttcaggaccttgacatgatg 1345
Db 1366 aagttgttgagattaccc-----agagaaacctcagcttcgggacttgatgtcttc 1419
QY 1346 ctcacctgtgtagatcaaaagaaggagctttgatgtctccacacttcaactcaaaaggcc 1405
Db 1420 ctcagtgttggatatgaacgaggggagctcttttctaccacacttcaattcaaaaggcc 1479
QY 1406 atggttatcgtcgtcgtcaacaaaggaaactggaaaccttgaactcgtggctgtaagaaaa 1465
Db 1480 atagtggtactagtattaatgaaggagaagcaaacatgaacttgt---tggcattaaa 1536
QY 1466 gagcaacaacagaggggacggcggggaagaagaggaggagcgaagagaggaggga 1525
Db 1537 gaacaacaacagaggcgagcaacaggaagagcaacctttggaa----- 1578
QY 1526 agtaacagagaggtgcgtaggtacacagcgaggttgaaagaggcgatgtgttcatcatg 1585

Db 1579 -----gtcggaatatagagctgaattgtctgaacaagatatattgtaatc 1626

QY 1586 ccagcagctcatccagtagccatcaacgcttccctccgaactccatctgcttggtcggt 1645

Db 1627 ccagcaggttatccagttatggtcaacgctacctcagatctgaatttcttgctttggt 1686

QY 1646 atcaacgctgaacaacacacagaaatcttcttgagggtgataaggacaaatgtatagac 1705

Db 1687 atcaatgccgagacaacacagaggaaacttcttgagggttcgaaagacaatgtataagc 1746

QY 1706 cagatagagaagcaagcgaaggatttagcattccctgggtcggtgaacaagttgagaag 1765

Db 1747 cagataccttagtcaagtgcaggagcttggttccctagggtcgtgcaaaagatataggaaac 1806

QY 1766 ctcatcaaaaaccagaaggaatctcactttgtgagtgctcgtctca 1812

Db 1807 ctaataaagagccaaagtgaagtcctactttgtggatgctcagcctca 1853

RESULT 10

AAV42316

ID AAV42316 standard; cDNA; 2140 BP.

XX AAV42316;

AC AAV42316;

XX 27-OCT-1998 (first entry)

DT

XX Macadamia integrifolia partialantimicrobial protein gene.

DE

XX antimicrobial protein; infestation; control; ss.

KW

XX Macadamia integrifolia.

OS

XX Key Location/Qualifiers

FT CDS 1..1878

FT /*tag= a

FT /product= partial antimicrobial protein

FT mat_peptide 1..1875

FT /*tag= b

FT

XX WO9827805-A1.

PN

XX 02-JUL-1998.

PD

XX 22-DEC-1997; 97WO-AU00874.

PF

XX 20-DEC-1996; 96AU-0004275.

PR

XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

PA

XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

PI

XX WPI; 1998-377279/32.

DR

DR P-PSDB; AAW62830.

XX

XX Novel anti-microbial protein from e.g. Macadamia integrifolia -

PT useful for controlling microbial infestations of plants or mammals

XX

PS Claim 5; Page 46-47; 96pp; English.

XX

CC The sequence is that encoding an antimicrobial protein which can

CC be used to control microbial infestations in plants and mammalian

CC animals.

XX

SQ Sequence 2140 BP; 695 A; 486 C; 557 G; 402 T; 0 other;

Query Match 7.9%; Score 160.2; DB 19; Length 2140;

Best Local Similarity 48.5%; Pred. No. 6.4e-37;

Matches 705; Conservative 0; Mismatches 658; Indels 90; Gaps 6;

QY 415 gcgtgaaagagaagagactggagacaaaccaagaagattggaggcgaccaagtcatca 474

Db 471 gcggcactgcgaacaacaggagccacgtctgcagtcaccagtgccagcgaaagatgccaaaga 530

QY 475 gcagccacgggaaaaataaaggccggaaggaaagagaagagaacaagagtggggaaacaccagg 534

Db 531 gcagcagaggcaacacacggcgaggtggcgatttgatgaaccctcagaggggagcgagcgg 590

QY 535 tag---ccatgtgaggggaagaacaatctcgggaacacccctttctacttcccgtaagggcg 591

Db 591 cagatacagaggaggagaagagaagcaaaagcgacaacccctactactctcagacgaacgaag 650

QY 592 gtttagcaccgcgtacggggaaccaaaccgggtaggatccgggtcctgcagaggtttgacca 651

Db 651 cttaagtacaaggttcaggaccgaggaaggccacatctcagttctggagaacttctatgg 710

QY 652 aaggtcaaggcagtttccagaatctccagaatccacogtatgtgcagatcgagggccaaacc 711

Db 711 tagatccaagcttctacgcgcactaaaaaactatcgcttggtgctcctcgaggctaacc 770

QY 712 taacactctgttcttcccaagcacgctgatgctgataaacatccctgttatccagcaagg 771

Db 771 caacgccttcgtgctccctaccacacttggtgcagatgccatctctcttggtcatcgagg 830

QY 772 gcaagccaccgtagaccgtagcaaaatggcaataacacagaagaagcttttaattcttgacgagg 831

Db 831 gagaggagccctcaaaatgatccaccgtgacaacacagagaatcctacaacctcgagtgtgg 890

QY 832 ccattgcactcagaatcccatccggtttcatttcttccatcttgaaacccgcacatgacaacca 891

Db 891 agacgtaaatcagaatcccgctggaaccacacattctacttaatacaaccgagacaacaaga 950

QY 892 gaacctcagagtagctaaaaatctccatgcccggttaacacacacccggccagtttgaggttt 951

Db 951 gaggctccacatagccaagttctttacagaccatactccactcctggccaatacaaggaatt 1010

QY 952 ctctccggcgagcagccgagaccaatcatctctacttgcagggttcagcagggaatacgtt 1011

Db 1011 ctctccagctggaggccaaaacccagagccgtacctcagttaccttcagcaaaagattct 1070

QY 1012 ggaggcgcttcaatgcgggaattcaatgagatacggagggtgctgttagaagagaatgc 1071

Db 1071 cgaggctgcgtcaacacacaaacagagaggtgcgtgggtgct----- 1115

QY 1072 aggaggtgagcaagaggagaggggcagaggcgatggagtactcggagtagtgagaacaa 1131

Db 1116 -----tggaacagcaaaag 1127

QY 1132 tgaaggagtagtagtcaaaagtgtcaaaaggagacgcttgaagaacttact---aagcacgc 1188

Db 1128 ggaggagtgataaattaggcgctcaccaggagcagatcaggaggttgactcgagatgactc 1187

QY 1189 taaatccgtctcaaaagaaaggctccgaagaagaagaggagagatatcaccaacccaatcaactt 1248

Db 1188 agagtcaacgacgctggcataataaggagaggtggtgaatcaacgaggggaccttacaatct 1247

QY 1249 gagagaaggcgagcccgatcttttaacaacacttttggaagttatttgaggtgaagccaga 1308

Db 1248 gttcaacaaaaaggccactgtacttccaacaaaatacgggtcaagcctacgaagtcaaacctg- 1306

QY 1309 caagaagaacccccagcttcaggacctggacatgatgctcacctgtgtagagatcaaaaga 1368

Db 1307 --aggactacaggaactccaagacatggacgtatcaagtttcatagccaacatcaccca 1364

QY 1369 aggagcttgatgctccccacacttcaactcaaaaggccatggttatcgtcgtcgtcaacaa 1428

Db 1365 gggatccatgatgggtcccttcttcaacactaggtctcaaaaggtggtggtggtgtag 1424

QY 1429 aggaactgaaacaccttgaactcgtggctgtaagaaaaagagaacaacagaggggacggcg 1488

Db 1425 tggagaggcagatgtggaatggcatgccctcacttgcgtcggaagacacgcgcgcgcg 1484

QY 1489 ggaagaagaggaggacgaagacgaagaagagaggagggaagttaacagagaggtgcgtaggt 1548

Db 1485 tggaggaaaaagcatgaggagggaagaggaggtgcactatgagcagg-----t 1532

QY 1549 cacagcaggttgaaggagcgcgtgtgttcacatgccagcagctcaccagtagccat 1608

Db 1533 tagagcagcttgcgaagagagagggccattgtgttctgaggtcaccgcgtcgtctt 1592

QY 1609 caacgcttcctccgaaactccatctgc-----ttggcttcggtatcaacgctgaaaaaa 1662

Db 1593 cgtttcaccggaacgaaaaacctgctgctttttgcatcttggaatcaatgcccaaaaaa 1652

QY 1663 ccacagaatcttcttcaggtgataaggacaaatgtgatagaccagatagagaagcaagc 1722

Db 1653 ccacgagaacttcctcgcggggagagagaggaaacgtgctgcagcagatagagccacagc 1712

QY 1723 gaaggatttagcattccctgggtcggtgaaacaagttgagaagctcatcaaaaaaccagaa 1782

Db 1713 aatggagctagcgtttgccgttcaaggaaagaggtagaagagttatttaacagccagga 1772

QY 1783 ggaatctcaactttgtgagtgctcctcctcaatctcaatctcaatctcctcgtctcctga 1842

Db 1773 cgagctcatcttcttctcctggccaggcagcaccagcaaacagtcgccccgcctccacaa 1832

QY 1843 gaaagagctctcct 1855

Db 1833 gcaacaacagcct 1845

RESULT 11

AAV42311

ID AAV42311 standard; cDNA; 2171 BP.

XX

AC AAV42311;

XX

DT 27-OCT-1998 (first entry)

XX

DE Macadamia integrifolia antimicrobial protein gene.

XX

KW antimicrobial protein; infestation; control; ss.

XX

OS Macadamia integrifolia.

Key Location/Qualifiers

FT CDS 1..2001

FT /tag= a

FT /product= antimicrobial protein

FT sig_peptide 1..86

FT /tag= b

FT mat_peptide 87..1999

FT /tag= c

XX WO9827805-A1.

XX

PN 02-JUL-1998.

XX

PD 22-DEC-1997; 97WO-AU00874.

XX

PR 20-DEC-1996; 96AU-0004275.

XX

XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

PA

XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;

PI WPI; 1998-377279/32.

XX P-PSDB; AAW62829.

DR

XX Novel anti-microbial protein from e.g. Macadamia integrifolia -

PT useful for controlling microbial infestations of plants or mammals

XX

PS Claim 5; Page 41-43; 96pp; English.

XX

CC The sequence is that encoding an antimicrobial protein which can

CC be used to control microbial infestations in plants and mammalian

CC animals.

XX

SQ Sequence 2171 BP; 680 A; 509 C; 571 G; 411 T; 0 other;

Query Match 7.7%; Score 157.4; DB 19; Length 2171;

Best Local Similarity 48.9%; Pred. No. 4.e-36;

Matches 649; Conservative 0; Mismatches 591; Indels 87; Gaps 5;

QY 538 ccatgtgagggaagaaacatctcggaaacacccctttctacttcccgtaagcggttttag 597

Db 720 cgaggaggagagaagcaaaagcgacaacccctacttctcgacgaacgaagcttaag 779

QY 598 caccgcgtacgggaaccaaacggtaggtatccgggtcctcgagaggttttgaccaaaaggtc 657

Db 780 tacaaggttcaggaccgaggaaggccacatctcagttcttgagaaacttctatggtagatc 839

QY 658 aaggcagtttcagaatctccagaatcacccgtattgtgcagatcgaggccaaacctaacac 717

Db 840 caagcttctacgcgcactaaaaaactatcgttggctcctcgaggctaaccaccaacgc 899

QY 718 tctgttcttccaaagcacgctgatctgataacatcctgttatccagcaagggcaagc 777

Db 900 cttcgtgctcctaccacattggacgcagatgccattctcttggtcaccggaggagaggg 959

QY 778 caccgtgaccgtagcaaatggcaataacagaaagagctttaatcttgacgagggccatgc 837

Db 960 agccctcaaatgatccaccgtgacaacagagaatcctacaacctcgagtggtgagacgt 1019

QY 838 actcagaatcccatccggttttcatttctacatcttgaaacccgcatgacaaccagaacct 897

Db 1020 aatcagaatcccgctggaaacacattctacttaatacaccgagacaacaacgagaggtc 1079

QY 898 cagagtagctaaaaatctccatgccgtttaacacacaccccgccaggtttgaggtttcttccc 957

Db 1080 ccacatagccaagtcttacagaccatatccactcctggccaatacaaggaattcttccc 1139

QY 958 ggcgagcagcagagaccaatcatcctacttgcagggttcagggttcagcaggaatactgtggaggc 1017

Db 1140 agctggaggccaaaacccagagccgtacctcagttaccttcagcaaaagagattctcgaggc 1199

QY 1018 cgccttcaatcggaatttcaatgagatcacggagggtgctgtttagaagagaatgcaggaggg 1077

Db 1200 tgcgtcaacacacaagcagagagggctgctgggtgct----- 1238

QY 1078 tgagcaagagagagaggggcagagggcgatggagtactcggagtagtgagaaacaatgaagg 1137

Db 1239 -----tggacagcaaaaggaggg 1256

QY 1138 agtगतatgtaaaagtgtcaaaaggagcagcttgaagaacttact---aagcacgctaaaatc 1194

Db 1257 agtgataattagtcggtcacagggagcagatcaggaggttgactcgagatgactcagagtc 1316

QY 1195 cgtctcaaaagaaaggctccgaagagaggagagatatcacaacccaactcaacttgagaga 1254

Db 1317 acgacgctggcatataaaggagaggtggtgaatcaagcagggggacattacaatctgttcaa 1376

QY 1255 aggcgagcccgatcttttcaacaactttgggaagtatttgaggtgaagccagacaagaa 1314

Db 1377 caaaaggccactgtactccaacaatacgggtcaagcctcacgaagtcaaacctg---agga 1433

QY 1315 gaacccccagcttcaggacccgtggacatgatgctcacctgtgtagatcaaaagaaggagc 1374

Db 1434 ctacaggccaactccaagacatggacgtatcggttttcataagccaacatcacccagggatc 1493

QY 1375 ttgtatgctccacacttcaactcaaaaggccatggttatcgtcgtcgtaacaaagggaac 1434

Db 1494 catgatgggtcccttcttcaacactaggtctacaaaagggtggttagtggtggttagtgaga 1553

QY 1435 tggaaaaccttgaactcgtggtgctgtaagaaaaagagcaacaacagaggggacggtgggaaga 1494

Db 1554 ggcagatgtggaaatggcatgccctcacttgtcgggaagacacgcggcgccgctggaggg 1613

QY 1495 agaggaggacgaagacgaagaaggagggaagtaacagagaggtgcgttaggtacacagc 1554

Db 1614 gaaaagcatgaggagagagatgtgcactatgag-----caggttaaagc 1661
QY 1555 gaggttgaaggagcgatgtttcatcatgcccagcagctcatccagtagccatcaacgc 1614
Db 1662 acgtttgtcgaagagagagccattgttgcggttaggtcatcccgctcgtcttcgtttc 1721
QY 1615 ttc-----ctccgaactccatctgcttggcttgcgttatcaacgctgaaaaacaaccacag 1668
Db 1722 atcggaaacgagaacctgctgttttgcatttggaatcaatgcaccaaaacaaccacga 1781
QY 1669 aatcttcttgccaggtgataaaggacaatgtgatagaccagatagagaagcaagcgaagga 1728
Db 1782 gaacttctcgcggtgagagaggaacgtgtgctgcagcagatagagccacaggaatgga 1841
QY 1729 tttagcattccctgggtggtggaacaaagttagaagctcatcaaaaaaccagaaggaatc 1788
Db 1842 gctagcgtttccgctccaaggaaagaggttagaagattatttaacagccaggacgagtc 1901
QY 1789 tcactttgtgagtgtcctcctcaatctcaatctcaatctcctcgtcgtctcctgagaaaaa 1848
Db 1902 tatcttcttctcgtggccaggcagcaccagcaacagtcttcccgctccaccaagcaaca 1961
QY 1849 gtctcct 1855
Db 1962 acagcct 1968

RESULT 12
AAV42310
ID AAV42310 standard; cDNA; 2171 BP.
XX AAV42310;
AC
XX
DT 27-OCT-1998 (first entry)
XX
DE Macadamia integrifolia antimicrobial protein gene.
XX
KW antimicrobial protein; infestation; control; ss.
XX
OS Macadamia integrifolia.
XX
FH Key Location/Qualifiers
FT CDS 1..2001
FT /*tag= a
FT /product= antimicrobial protein
FT sig_peptide 1..85
FT /*tag= b
FT mat_peptide 86..1999
FT /*tag= c

XX
PN WO9827805-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-AU00874.
XX
PR 20-DEC-1996; 96AU-0004275.
XX
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR WPI; 1998-377279/32.
DR P-PSDB; AAW62828.
XX
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
PS Claim 5; Page 37-38; 96pp; English.
XX
CC The sequence is that encoding an antimicrobial protein which can
be used to control microbial infestations in plants and mammalian

CC animals.
XX
SQ Sequence 2171 BP; 687 A; 507 C; 570 G; 407 T; 0 other;

Query Match 7.5%; Score 152.6; DB 19; Length 2171;
Best Local Similarity 48.7%; Pred. No. 1.2e-34;
Matches 646; Conservative 0; Mismatches 594; Indels 87; Gaps 5;

QY 538 ccattgtgagggaagaacatctcgaacaaccccttttacttcccgctcaaggcggttag 597
Db 720 cgaggagggaagaggagcaaaagcgacaacccctactacttcgacgaacgaagcttaag 779
QY 598 caccgcgtacggggaaccaaaccggttaggatccgggtccttcgagaggtttgaccaaagtc 657
Db 780 tacaaggttcaggaccgagggaagccacatctcagttctggagaacttctatgtagatc 839
QY 658 aaggcagtttcagaatctccagaatcaccgtattgtgcagatcgagggcacaacctaaccac 717
Db 840 caagcttctacgcgcactaaaaaactatcgcttgggtcctcgcaggttaacccaacgc 899
QY 718 tcttgttcttcccaagcacgctgctgataaacatccttctgttatccagcaaggccaagc 777
Db 900 ctctcgtcctcctacccttggatgcagatgccattctcttggtcattagggagagag 959
QY 778 caccgtgaccgtagcaaatggcaataacagaaagagctttaatcttgacgagggccatgc 837
Db 960 agccctcaaatgatccaccacgacaacagagaatcctacaacctcgagtgtggagacgt 1019
QY 838 actcagaatcccatccggttttcttctctacatcttgaaccgcatgacaacccaacac 897
Db 1020 aatcagaatcccgctggaaaccattctacttaataacacgagacaacaacagagagct 1079
QY 898 cagagtgtctaaaaatctccatgcccggttaacacacaccccgccagtttgaggatttctccc 957
Db 1080 ccacatagccaagtctttacagaccatatccactcctggccaatacaaggaaattctccc 1139
QY 958 ggcgagcagccgagaccaatcatcctactcttcagggcttcagcaggaatacgttggaggc 1017
Db 1140 agctggaggccaacacccagagccgtacctcagtaaccttcagcaaaagattctcgaggc 1199
QY 1018 cgccttcaatgcgggaattcaatgagatcagggaggtgctgttagaagagaatgcaggag 1077
Db 1200 tgcgtcaacacacaaacagagaagctgcgtgggtgtt----- 1238
QY 1078 tgagcaaggagagagggcagagggcagtgagtgagtagtgagaacaatgaagg 1137
Db 1239 -----tggaacagcaaggaggag 1256
QY 1138 agtgatgtcaaatgttcaaaaggagcacgttgaagaacttact---aagcacgtataatc 1194
Db 1257 agtgataattagggcgtcacaggagcagatcagggagttgactcgagatgactcagagtc 1316
QY 1195 cgtctcaaaagaggctccgaagaagaggagagatatcaccaacccaatcaacttgagaga 1254
Db 1317 acgacactggcatataaggagaggtgtggaatcaagcaggggaccttacaatctgttcaa 1376
QY 1255 agcgagcccgatctttctaaacactttgggaagtattttgaggtgaagccagacaagaa 1314
Db 1377 caaaagccactgtactccaacaataacggtcaagccctacgaagtcaaacctg---agga 1433
QY 1315 gaacccccagcttcaggacctggacatgatgctcacctgtgtagagatcaaaagaaggagc 1374
Db 1434 ctacaggcaactccaagacatggacttatcggttttcatagccaacgtcaccaggagtc 1493
QY 1375 tttagtctccacacttcaactcaaaaggcccatggtttatcgtcgtcgtcaacaagaagaa 1434
Db 1494 catgatgggtcccttcttcaacactaggtctacaaaggtggtagtggtgtagtgaga 1553
QY 1435 tggaaaaccttgaaactcgtggtgttaagaaaaagagcaacaacagaggggacggcggaaga 1494
Db 1554 ggcagatgtggaaatggcatgccctcacttctcgggaagacacggcgccggtggagag 1613

Db 1248 gatatgggtacgctcaaatggcttgcccgcatctctccagacagagccaggatcccaaa 1307
QY 1479 ggggacggcgggaagagagaggagacgaacgaagaagagagggaagtaacagagagg 1538
Db 1308 gtggaaggcaagacagaagaaacaagaagaagatcagaagaaggagacatttggaagaat 1367
QY 1539 tgcgtaggtacacagcagaggttgaaggaaggcgcgatgtgttcacatgccagcagctcatc 1598
Db 1368 tccagcaggtcaaagccccattgtcacctgggtgacgtctttgtagccccgcaggccatg 1427
QY 1599 cagtagccatcaacgcttccctccgaactccatctg-----cttggcttcggtatcaacg 1652
Db 1428 cagttacattctttgcatccaaagaccagccctgaatgcagttgctgttggactcaacg 1487
QY 1653 ctgaaaaacaaccacagaatcttcccttgcaaggtgataaaggacaaatgtgatagaccagatag 1712
Db 1488 ccagaacaaccagagaattttcccttgcaagggaaaaaagaacttggtcaga---caaatgg 1544
QY 1713 agaagcaagcgaagatttagcattccctgggtcggtgaacaagttgagaagctcatca 1772
Db 1545 atagcgaggcaaaaggagttatcatttgggtaccatcgaaattggtagataatatattca 1604
QY 1773 aaaaccagaaggaatctcaactttgtg 1798
Db 1605 acaaccggatgagtcgtatttcatg 1630

RESULT 14
AAV72243
ID AAV72243 standard; cDNA; 1924 BP.
XX
AC AAV72243;
XX
DT 24-MAY-1999 (first entry)
XX
DE G. max SBP2 cDNA.
XX
KW SBP1; sucrose binding protein; SBP2; sucrose uptake; transgenic plant;
KW seed; carbohydrate content; soybean; ds.
OS Glycine max.

Key Location/Qualifiers
CDS 119..1588
FT /*tag= a
FT /product= "SBP2"

PN WO9853086-A1.
XX
PD 26-NOV-1998.
XX
PF 21-MAY-1998; 98WO-US10465.
XX
PR 22-MAY-1997; 97US-0047568.
XX

(UNIW) UNIV WASHINGTON STATE RES FOUND.
PA Chao WS, Grimes HD;
PI WPI; 1999-070155/06.
XX P-PSDB; AAW90341.
DR
XX

PT New modified plant sucrose binding proteins - used to develop
PT transgenic plants which can have enhanced or decreased sucrose
PT uptake activity in developing seeds
XX
PS Disclosure; Page 40-42; 58pp; English.
XX

CC This sequence encodes a novel sucrose binding protein, SBP2 isolated
CC from Glycine max. The SBP2 protein is used in a method resulting in the
CC production of a modified plant sucrose binding protein (SBP) which has a
CC modified amino acid sequence compared to a corresponding wild-type SBP,
CC and where expression of the modified SBP in a yeast assay system confers

CC enhanced sucrose uptake compared to the corresponding wild-type SBP.
CC The products of the invention can be used for producing transgenic plants
CC which have modified sucrose uptake activity, particularly in developing
CC seeds. Enhanced sucrose uptake activity in developing seeds may be
CC desirable where it is an advantage to increase the carbohydrate content
CC of the seed (e.g. where the seed is the primary plant material harvested,
CC such as soybean). In contrast, decreased sucrose uptake activity in
CC seeds might be desirable where the vegetative material of the plant is
CC harvested. The SBP regulatory regions confer specific or enhanced
CC expression in developing seeds and so may be used to express any
XX transgene in developing seeds.
SQ Sequence 1924 BP; 548 A; 432 C; 471 G; 473 T; 0 other;

Query Match 4.9%; Score 98.8; DB 20; Length 1924;
Best Local Similarity 50.2%; Pred. No. 1e-18;
Matches 244; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

QY 593 tttagcaccgcgtacgggaaccaaaacggtaggtccggtcctgcagaggtttgaccaa 652
Db 437 ttttcgaccagatcgaaacagaaggtggcagcattcgggtctcgaagaagttcactgag 496
QY 653 aggtcaaggcagtttcagaaatctccagaatcacccgtattgtgcagatcgaggccaaacct 712
Db 497 aaatcccaagcttcttcaaggcattgagaatttccgtttggccattcttagaagctagaca 556
QY 713 aacactcttgttcttcccaagcgcgtgatgctgataacatccttgttatccagcaaggg 772
Db 557 cacacgttcgtgtccccacgcacttttgattccgaggttgcgttcttgcacaacataagg 616
QY 773 caagccaccgtgaccgtagcaaatggcaataaacagaaagagcctttaaactcttgacaggggc 832
Db 617 agagccgtacttgggttggtgagggaagtgaaacagaaaaaatcacccctagaacctgga 676
QY 833 catgcactcagaatccccatccggttttcatttctcattcctgaaccgcctatgacaaccag 892
Db 677 gacatgatacacataccagcagcgacacaccactgtacatcgtttaacagagatgagaatgag 736
QY 893 aacctcagagtagctaaaaatctccatgcccgtttaacacaccccggccagtttgaggatttc 952
Db 737 aagctcctcttgccatgctccatataacctgtctctactccttgaaaaatttgaggaaattt 796
QY 953 ttcccggcagcagccgagacccaatcactcacttgcagggttcacagcaggaatacgttg 1012
Db 797 ttcgggcctggaggacgagaccccaagaatcggtcctctcagcattcagctgggaatgtgctg 856
QY 1013 gaggccgccttcaatgcggaattcaatgagatacggagggtgctgttagaagagaatgca 1072
Db 857 caagctgcgtctccaaaccccccaaggaaagttagaagggtttttaatcaacagaacgag 916
QY 1073 ggaggt 1078
Db 917 ggaagt 922

RESULT 15
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
OS Synthetic.
XX
PN WO200107665-A2.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2001, 13:04:45 ; Search time 70.16 Seconds
(without alignments)
218.546 Million cell updates/sec

Title: US-09-715-036-2
Perfect score: 1114
Sequence: 1 MAKLTILVALFLAAHAS.....LPRAADGDGWFPVAVDCSG 207

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 1114 | 100.0 | 207 | 22 AAB82383 | Peanut allergen Ar |
| 2 | 824 | 74.0 | 156 | 20 AAY40973 | Ara h 2 protein fr |
| 3 | 824 | 74.0 | 157 | 18 AAW24164 | Peanut allergen Ar |
| 4 | 824 | 74.0 | 157 | 18 AAW24153 | Peanut allergen Ar |
| 5 | 824 | 74.0 | 157 | 20 AAY15245 | Peanut allergen, A |
| 6 | 824 | 74.0 | 157 | 22 AAU04707 | Anaphylactic anti |
| 7 | 753 | 67.6 | 166 | 20 AAY40968 | Recombinant Ara h |
| 8 | 683 | 61.3 | 167 | 22 AAU04710 | Modified anaphylac |
| 9 | 681 | 61.1 | 166 | 21 AAB33600 | Modified Ara h 2 a |
| 10 | 681 | 61.1 | 166 | 22 AAU05035 | Modified anaphylac |
| 11 | 253 | 22.7 | 158 | 18 AAW23419 | Soybean albumin 3 |

| | | | | | | |
|----|-------|------|-----|----|-----------|--------------------|
| 12 | 253 | 22.7 | 158 | 20 | AAAY05723 | Soybean Gm2S-1 sul |
| 13 | 251.5 | 22.6 | 155 | 18 | AAW23418 | Soybean albumin 1 |
| 14 | 239 | 21.5 | 158 | 18 | AAW23420 | Chimeric Soybean a |
| 15 | 162.5 | 14.6 | 158 | 18 | AAW23586 | Mabinlin MBLI from |
| 16 | 157.5 | 14.1 | 158 | 18 | AAW23588 | Mabinlin MBLIII fr |
| 17 | 156 | 14.0 | 28 | 18 | AAW24194 | Peanut allergen Ar |
| 18 | 150.5 | 13.5 | 155 | 18 | AAW10569 | Mabinlin II prepro |
| 19 | 150.5 | 13.5 | 155 | 18 | AAW23587 | Mabinlin MBLII fro |
| 20 | 147.5 | 13.2 | 164 | 10 | AAAP91892 | Amino acid sequenc |
| 21 | 147.5 | 13.2 | 164 | 10 | AAAP96144 | Sequence encoded b |
| 22 | 142 | 12.7 | 184 | 12 | AAR14193 | Napin 1-2 gene pro |
| 23 | 142 | 12.7 | 184 | 15 | AAR54981 | Napin protein from |
| 24 | 142 | 12.7 | 184 | 16 | AAAR76694 | B. campestris napi |
| 25 | 142 | 12.7 | 184 | 18 | AAW15415 | Napin encoded by c |
| 26 | 142 | 12.7 | 184 | 19 | AAW30526 | Brassica campestri |
| 27 | 142 | 12.7 | 184 | 21 | AAAY49926 | Brassica campestri |
| 28 | 142 | 12.7 | 189 | 9 | AAAP81140 | Sequence encoded b |
| 29 | 140 | 12.6 | 140 | 10 | AAAP91891 | Part of the sequen |
| 30 | 138 | 12.4 | 28 | 18 | AAW24199 | Peanut allergen Ar |
| 31 | 137 | 12.3 | 184 | 19 | AAW30670 | Napin seed storage |
| 32 | 137 | 12.3 | 189 | 15 | AAR54980 | Napin protein from |
| 33 | 135 | 12.1 | 171 | 15 | AAR53579 | Synthetic 2S seed |
| 34 | 129 | 11.6 | 110 | 15 | AAAR62762 | Synthetic 12 kd pr |
| 35 | 128 | 11.5 | 111 | 19 | AAW86264 | Wild-type Brazil n |
| 36 | 128 | 11.5 | 149 | 19 | AAW86272 | Modified Brazil nu |
| 37 | 128 | 11.5 | 149 | 19 | AAW86273 | Modified Brazil nu |
| 38 | 127 | 11.4 | 22 | 21 | AAB27527 | Ara h 2 residues 5 |
| 39 | 125.5 | 11.3 | 142 | 21 | AAAY84356 | Amino acid sequenc |
| 40 | 124 | 11.1 | 149 | 19 | AAW86274 | Modified Brazil nu |
| 41 | 120.5 | 10.8 | 174 | 22 | AAAY72901 | Flax 2S storage pr |
| 42 | 119 | 10.7 | 20 | 21 | AAB23115 | Peanut wild-type A |
| 43 | 119 | 10.7 | 20 | 21 | AAAB33564 | 5/20 native Ara h |
| 44 | 119 | 10.7 | 20 | 22 | AAU04777 | Native anaphylacti |
| 45 | 119 | 10.7 | 20 | 22 | AAU05102 | Anaphylactic peanu |

ALIGNMENTS

| | |
|----------|--|
| RESULT | 1 |
| AAB82383 | |
| ID | AAB82383 standard; Protein; 207 AA. |
| XX | |
| AC | AAB82383; |
| XX | |
| DT | 23-JUL-2001 (first entry) |
| XX | |
| DE | Peanut allergen Ara h2 gene product. |
| XX | |
| KW | Peanut; allergen; Ara h2; transgenic plant; allergy. |
| XX | |
| OS | Arachis hypogaea. |
| XX | |
| FH | Key |
| FT | Peptide |
| FT | Protein |
| FT | Location/Qualifiers |
| FT | 1..21 |
| FT | /label= Signal_peptide |
| FT | 22..207 |
| FT | /label= Mature_protein |
| XX | |
| PN | WO200136621-A2. |
| XX | |
| PD | 25-MAY-2001. |
| XX | |
| PF | 20-NOV-2000; 2000WO-US31657. |
| XX | |
| PR | 19-NOV-1999; 99US-0167255. |
| XX | |
| PA | (UYAL-) UNIV ALABAMA A & M. |
| XX | |
| PI | Dodo HW, Arntzen CJ, Konan KN, Viquez OM; |
| XX | |
| DR | WPI; 2001-355630/37. |
| DR | N-PSDB; AAF90336, AAF90337. |

XX Producing transgenic peanut plants that produce allergen-free seeds,
PT useful in non-allergenic foods, by antisense or sense co-suppression of
PT allergen-encoding genes -
XX
PS Example 1; Fig 2; 72pp; English.
XX
CC The present sequence is that of the peanut allergenic protein (AP)
CC encoded by the Ara h2 gene (see AAF90336). The invention relates
CC to a method for producing a peanut plant having reduced, or
CC undetectable, AP content in its seed. A peanut plant cell is
CC transformed with a DNA construct containing an antisense AP gene
CC and/or sense AP gene, or their fragments, regenerated to plants,
CC and fertile transgenic plants that produce seeds with reduced AP
CC content are identified. The AP sense or antisense gene may
CC comprise at least a portion of the Ara h2 gene sequence. The seeds
CC are useful for preparation of allergen-free foods. Recombinant AP
CC may be produced and used to produce antibodies useful for detecting
CC AP in foods, and for treatment or prevention of peanut allergy.
XX
SQ Sequence 207 AA;

Query Match 100.0%; Score 1114; DB 22; Length 207;
Best Local Similarity 100.0%; Pred. No. 4.8e-110;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKLTILVALALFLAAHASARQQWELQGRRCQSLERANLRPCEQHLMKIQRDEDSY 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 makltiilvalalflaaahasargqwelqgdrccqqlerlanlrpceqhlmqigrdedsy 60

QY 61 ERDPYSPSPYSPSPYDRRGAGSSQHQERCCNELNEFENNRCMCCEALQQIMENQSDRL 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 erdpypspqypspypdyrrgagssqhgerccnelnefennqrcmcealqgimengsdr1 120

QY 121 QGRQEQQFKRELNLPPQCGGLRAPQRCDDLVEGGRPRIPPILTGSRRRHQSPYGNR 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 qgrqeqqfkelrnlppqcgglrapqrcddlvesgrrprippiltsrsrhqsypgnr 180

QY 181 RYSAMCLLPRAADGDGWFPSVAVDCSG 207
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 rysamcllpraadgdgwfpsvavdcsg 207

RESULT 2
AA40973
ID AAY40973 standard; protein; 156 AA.
XX
AC AAY40973
XX
DT 06-DEC-1999 (first entry)
XX
DE Ara h 2 protein fragment.
XX
KW Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW allergic reaction; Ara h 2.
XX
OS Arachis hypogaea.
XX
PN WO9945961-A1.
XX
PD 16-SEP-1999.
XX
PF 12-MAR-1999; 99WO-US05494.
XX
PR 12-MAR-1998; 98US-0077763.
XX
PR 11-MAR-1999; 99US-0077763.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX

DR WPI; 1999-551218/46.
XX
PT Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction -
XX
PS Disclosure; Page 104; 193pp; English.
XX
CC The invention provides a tertiary structure for the peanut allergen
CC Ara H 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to
CC ensure that the allergen is not introduced into genetically modified
CC food. The present sequence represents an Ara h 2 protein fragment.
XX
SQ Sequence 156 AA;

Query Match 74.0%; Score 824; DB 20; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.8e-79;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LTILVALALFLAAHASARQQWELQGRRCQSLERANLRPCEQHLMKIQRDEDSYERD 63
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 lttilvalalflaaahasargqwelqgdrccqqlerlanlrpceqhlmqigrdedsyerd 60

QY 64 PYSQSPQYPSPYDRRGAGSSQHQERCCNELNEFENNRCMCCEALQQIMENQSDRLQGR 123
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 pypsqdpypspypdyrrgagssqhgerccnelnefennqrcmcealqgimengsdr1qgr 120

QY 124 QEQEQFKRELNLPPQCGGLRAPQRCDDLVEGGR 157
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 qeqqfkelrnlppqcgglrapqrcddlvesggr 154

RESULT 3
AAW24164
ID AAW24164 standard; Protein; 157 AA.
XX
AC AAW24164;
XX
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hII.
XX
KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy;
KW monoclonal antibody; ELISA; analysis; Ara hII.
XX
OS Arachis hypogaea strain Florunner.
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US15222.
XX
PR 04-MAR-1996; 96US-0610424.
PR 29-DEC-1995; 95US-0009455.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
XX
XX WPI; 1997-363453/33.
DR N-PSDB; AAT76614.
DR
XX
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in
XX two-site monoclonal antibody based ELISA
PS Claim 31; Page 219; 354pp; English.

XX CC This polypeptide comprises major peanut allergen Ara hII.
CC Its sequence was deduced from cDNA clone P38 (AAT76614), isolated
CC from peanut seed cDNA using a primer (see AAT76617) based on an
CC isolated Ara hI peptide (see AAW24151). The sequence shows
CC significant homology with the conglutin family of seed storage
CC proteins of other legumes. The allergen is recognised by serum
CC IgE from a large proportion of individuals with peanut
CC hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used to
CC raise monoclonal antibodies which are used in a specific two-site
CC MAb ELISA for the detection of Ara hI or Ara hII (claimed). IgE-
CC binding Ara hII antigen epitopes (see AAW24188-93) may be used in
CC vaccines to protect against allergic reactions to peanut allergens,
CC e.g. anaphylactic shock.
XX SQ Sequence 157 AA;

Query Match 74.0%; Score 824; DB 18; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.8e-79;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LTILVALALFLAAHASARQQWELQGRRCQSLERANLRPCEQHLMQIQDEDSYERD 63
Db 1 ltllvalalflaahasarqqwelqgrrcqqleranlrpceqhlmqigrdedsyerd 60
QY 64 PYSPSQDPYSPYDRRGAGSSQHQRCCNELNEFENNQRMCCEALQQIMENQSDRLQGR 123
Db 61 pypsqdpypspydrrgagssqhqrccnelnefennqrcmcealqqimengsdrllqgr 120
QY 124 QEQQFKRELRLNPQCGLRAPQRCDLDVESGGR 157
Db 121 qgeqgfkrelrlnpqcgglrapqrcdldvesggr 154

RESULT 4
AAW24153
ID AAW24153 standard; Protein; 157 AA.
XX AAW24153.
AC (first entry)
DT 29-DEC-1997
DE Peanut allergen Ara hII.
XX Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy;
KW monoclonal antibody; ELISA; analysis; Ara hII.
XX Arachis hypogaea strain Florunner.
OS WO9724139-A1.
XX 10-JUL-1997.
PD 23-SEP-1996; 96WO-US15222.
XX 04-MAR-1996; 96US-0610424.
PR 29-DEC-1995; 95US-0009455.
XX (UYAR-) UNIV ARKANSAS.
PA Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
PI WPI; 1997-363453/33.
XX N-PSDB; AAT76615.
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in
PT two-site monoclonal antibody based-ELISA
XX Claim 31; Page 198; 354pp; English.
PS This polypeptide comprises major peanut allergen Ara hII.
XX CC

CC Its sequence was deduced from a cDNA clone (AAT76615) isolated
CC from peanut seed cDNA using a primer (see AAT76617) based on an
CC isolated Ara hI peptide (see AAW24151). The sequence shows
CC significant homology with the conglutin family of seed storage
CC proteins of other legumes. The allergen is recognised by serum
CC IgE from a large proportion of individuals with peanut
CC hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used to
CC raise monoclonal antibodies which are used in a specific two-site
CC MAb ELISA for the detection of Ara hI or Ara hII (claimed). IgE-
CC binding Ara hII antigen epitopes (see AAW24188-93) may be used in
CC vaccines to protect against allergic reactions to peanut allergens,
CC e.g. anaphylactic shock.
XX SQ Sequence 157 AA;

Query Match 74.0%; Score 824; DB 18; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.8e-79;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LTILVALALFLAAHASARQQWELQGRRCQSLERANLRPCEQHLMQIQDEDSYERD 63
Db 1 ltllvalalflaahasarqqwelqgrrcqqleranlrpceqhlmqigrdedsyerd 60
QY 64 PYSPSQDPYSPYDRRGAGSSQHQRCCNELNEFENNQRMCCEALQQIMENQSDRLQGR 123
Db 61 pypsqdpypspydrrgagssqhqrccnelnefennqrcmcealqqimengsdrllqgr 120
QY 124 QEQQFKRELRLNPQCGLRAPQRCDLDVESGGR 157
Db 121 qgeqgfkrelrlnpqcgglrapqrcdldvesggr 154

RESULT 5
AAW15245
ID AAW15245 standard; Protein; 157 AA.
XX AAW15245.
AC (first entry)
DT 09-NOV-1999
DE Peanut allergen, Ara h 2, amino acid sequence.
XX allergy; immune response; transgenic; allergen; epitope;
KW immunoglobulin E; Ig E; binding site; peanut.
XX Arachis hypogaea.
OS WO9938978-A1.
XX 05-AUG-1999.
PD 29-JAN-1999; 99WO-US02031.
XX 27-AUG-1998; 98US-0141220.
PR 31-JAN-1998; 98US-0073283.
PR 13-FEB-1998; 98US-0074590.
PR 13-FEB-1998; 98US-0074624.
PR 13-FEB-1998; 98US-0074633.
XX (SOSI/) SOSIN H.
PA (UYAR-) UNIV ARKANSAS.
PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
XX Bannon GA, Burks AW, Sampson HA, Sosin H;
PI WPI; 1999-479189/40.
XX N-PSDB; AAZ06383.
PT Modified allergen with reduced IgE binding, useful for treating e.g.
PT allergies
XX Disclosure; Page 38; 46pp; English.
PS

XX This is the amino acid sequence of the Ara h 2 protein from Arachis
CC hypogaea. The Ara h 2 protein has 10 IgE (Immunoglobulin E) binding
CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and
CC AAY15276).
CC By modifying the IgE binding sites the ability of the allergen to
CC provoke an immune response is downregulated. The epitopes of the IgE
CC binding sites can therefore be modified in genetically engineered plants
CC and animals to elicit less of an allergic response.
XX Sequence 157 AA;
SQ
Query Match 74.0%; Score 824; DB 20; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.8e-79;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 LTILVALALFLAAHASARQWELQGDRCQSLERANLRPCEQHLQMOKIQRDEDSYERD 63
Db 1 ltilvalalflaahasarqwwelqgdrrcqsleranlrpceqhlmqkigrdedsyerd 60
QY 64 PYSPODPYSPYDRRGAGSSQHQRCCNELNEFENNRCMCCEALQQIMENQSDRLQGR 123
Db 61 pypsqdpypspypdyrrgagssqhqrccnelnefennqrcmcealqqimengsdrlqgr 120
QY 124 QQEQQFKRELRLNPQQCGLRAPORCDLDVESGGR 157
Db 121 qgeqgfkrelrlnpqqcglrapgrcdldvesggr 154
RESULT 6
AAU04707
ID AAU04707 standard; Protein; 157 AA.
XX AAU04707;
XX 23-OCT-2001 (first entry)
XX Anaphylactic antigen Ara h 2.
DE
XX Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse.
XX Mus sp.
OS
XX WO200140264-A2.
PN
XX 07-JUN-2001.
PD
XX 06-DEC-2000; 2000WO-US33124.
PF
XX 06-DEC-1999; 99US-0455294.
PR
XX 23-JUN-2000; 2000US-0213765.
PR
XX 27-SEP-2000; 2000US-0235797.
XX
PA (PANA-) PANACEA PHARM LLC.
PA (UYAR-) UNIV ARKANSAS.
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX
PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
DR
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
XX the severity and/or number of allergic symptoms in individuals
PT sensitive to antigens, have reduced ability to bind Immunoglobulin E -
PT
XX Claim 7; Fig 10; 100pp; English.
PS
XX The sequence represents the amino acid sequence of anaphylactic antigen
CC Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design
CC antigenic peptides having a reduced ability to bind immunoglobulin E
CC (IgE) as compared with the intact (A), or having a sequence substantially

CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the
CC antigen, which is selected, formulated, and delivered so that binding of
CC the peptide to antigen-specific IgE is reduced as compared with IgE
CC binding of intact antigen. The composition is also useful for treating
CC and preventing allergic reactions.
XX Sequence 157 AA;
SQ
Query Match 74.0%; Score 824; DB 22; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.8e-79;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 LTILVALALFLAAHASARQWELQGDRCQSLERANLRPCEQHLQMOKIQRDEDSYERD 63
Db 1 ltilvalalflaahasarqwwelqgdrrcqsleranlrpceqhlmqkigrdedsyerd 60
QY 64 PYSPODPYSPYDRRGAGSSQHQRCCNELNEFENNRCMCCEALQQIMENQSDRLQGR 123
Db 61 pypsqdpypspypdyrrgagssqhqrccnelnefennqrcmcealqqimengsdrlqgr 120
QY 124 QQEQQFKRELRLNPQQCGLRAPORCDLDVESGGR 157
Db 121 qgeqgfkrelrlnpqqcglrapgrcdldvesggr 154
RESULT 7
AAAY40968
ID AAAY40968 standard; protein; 166 AA.
XX AAAY40968;
XX 06-DEC-1999 (first entry)
XX Recombinant Ara h 2 protein sequence.
DE
XX Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW allergic reaction; Ara h 2.
KW
XX Arachis hypogea.
OS
XX WO9945961-A1.
PN
XX 16-SEP-1999.
PD
XX 12-MAR-1999; 99WO-US05494.
PF
XX 12-MAR-1998; 98US-0077763.
PR
XX 11-MAR-1999; 99US-0077763.
PR
XX (UYAR-) UNIV ARKANSAS.
PA
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
PI
XX WPI; 1999-551218/46.
DR
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction -
PT
XX Disclosure; Page 81; 193pp; English.
PS
XX The invention provides a tertiary structure for the peanut allergen
CC

CC Ara H 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to
CC ensure that the allergen is not introduced into genetically modified
CC food. The present sequence represents a recombinant Ara h 2 protein.
XX
SQ Sequence 166 AA;

Query Match 67.6%; Score 753; DB 20; Length 166;
Best Local Similarity 100.0%; Pred. No. 6.5e-72;
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SARQQWELQGRRCQSLERANLRPCQHLQMKIQDEDSYERDPYSPSPYDR 79
Db 15 sarqqwelqgdrccqsleranlrpceqhlmqkqrddedsyerdypspqdpyspydr 74
QY 80 RGAGSSQHQRCCNELNEFENNRCMCALQQIMENQSDRLQGRQEQQFKRELRLNPQQ 139
Db 75 rgagssqhgerccnelnefennqrcmcealqgimengsdrlqgrqeqqfkrelrlnpqq 134
QY 140 CGLRAPQRCDDLVDVSGGR 157
Db 135 cglrapqrddldvesggr 152

RESULT 8
AAU04710
ID AAU04710 standard; Protein; 167 AA.
XX
AC AAU04710;
XX
DT 23-OCT-2001 (first entry)
XX
DE Modified anaphylactic antigen Ara h 2.
XX
KW Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO200140264-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US33124.
XX
PR 06-DEC-1999; 99US-0455294.
PR 23-JUN-2000; 2000US-0213765.
PR 27-SEP-2000; 2000US-0235797.
XX
PA (PANA-) PANACEA PHARM LLC.
PA (UYAR-) UNIV ARKANSAS.
XX
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX
PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX
DR WPI; 2001-381378/40.
XX

Antigenic fragments useful for reducing anaphylactic risk and reducing
PT the severity and/or number of allergic symptoms in individuals
PT sensitive to antigens, have reduced ability to bind Immunoglobulin E -
XX
PS Disclosure; Fig 12; 100pp; English.

XX
CC The sequence represents the amino acid sequence of modified anaphylactic
CC antigen Ara h 2, which has been altered to disrupt immunoglobulin E (IgE)
CC binding sites. Ara h 2 is an anaphylactic antigen (A), which was used to
CC design antigenic peptides having a reduced ability to bind IgE as

CC compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the
CC antigen, which is selected, formulated, and delivered so that binding of
CC the peptide to antigen-specific IgE is reduced as compared with IgE
CC binding of intact antigen. The composition is also useful for treating
CC and preventing allergic reactions.
XX
SQ Sequence 167 AA;

Query Match 61.3%; Score 683; DB 22; Length 167;
Best Local Similarity 91.5%; Pred. No. 1.7e-64;
Matches 129; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 20 SARQQWELQGRRCQSLERANLRPCQHLQMKIQDEDSYERDPYSPSPYDR 79
Db 15 sarqqaelqgdrccqsleranlraceahlmqkqadedsyerapysqapyspydr 74
QY 80 RGAGSSQHQRCCNELNEFENNRCMCALQQIMENQSDRLQGRQEQQFKRELRLNPQQ 139
Db 75 rgagssqhgerccnelnefennqrcmcealqgimengsdrlqgaqgqgkcrearnlpqq 134
QY 140 CGLRAPQRCDDLVDVSGGRRPR 160
Db 135 cglrapqrddadvesggrddr 155

RESULT 9
AAB33600
ID AAB33600 standard; Protein; 166 AA.
XX
AC AAB33600;
XX
DT 22-JAN-2001 (first entry)
XX
DE Modified Ara h 2 amino acid sequence.
XX
KW Allergy; allergic reaction; allergen; anaphylactic antigen; peanut;
KW Ara h 1; Ara h 2; Ara h 3; epitope; binding; immunoglobulin E; IgE;
KW food antigen; sensitising; immune response; anti-allergic.
XX
OS Arachis hypogea.
XX
PN WO200051647-A2.
XX
PD 08-SEP-2000.
XX
PF 03-MAR-2000; 2000WO-US05655.
XX
PR 03-MAR-1999; 99US-0122960.
PR 06-DEC-1999; 99US-0455294.
XX
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.
XX
PI Sampson HA;
XX
DR WPI; 2000-611341/58.
XX

PT Non-human animal sensitized to an antigen, useful as an animal model
PT for studying allergic reactions to allergens, such as those in food and
PT in the environment -
XX
PS Example 6; Fig 17B; 124pp; English.

XX The present invention describes an animal model which can be used for
CC studying allergic reactions to allergens. The animal is sensitised to
CC a selected antigen by administering the antigen itself or a nucleic
CC acid encoding the antigen, where preferably the antigen is an
CC anaphylactic antigen. The sensitised animal can then be used to screen
CC for compounds which may help to prevent, ameliorate, or cure allergic
CC conditions in humans. The animal model can be used for studying allergic
CC reactions to allergens, such as those in food (peanuts, fruit, berry,
CC nut, bean, milk, dairy products), or in the environment (weed pollen,
CC grass pollen, tree pollen, mite, animal, animal dander, fungal, and
CC insect antigens). AAB33478 to AAB33601 represent sequences which are
CC used in examples from the present invention to specifically examine the
CC peanut allergy, and the peanut anaphylactic antigens Ara h 1 to 3.
XX
SQ Sequence 166 AA;

Query Match 61.1%; Score 681; DB 21; Length 166;
Best Local Similarity 92.8%; Pred. No. 2.8e-64;
Matches 128; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 SARQQWELQGDRCQSLERANLRPCQHLQMQIKQDEDSYERDPYSPSPYDR 79
||||| ||||||| ||||| || ||||||| ||||||| ||||||| ||||||| |||||||
Db 15 sarqqaqlgddrrcqsqqlaranlraceahlmqkqadedsyerapyspyspydr 74

QY 80 RGAGSSQHQRCCNELNEFENNQCRCMCEALQQIMENQSDRLQGRQQEQFKRELRLNPQQ 139
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 75 rgagssqhqrccnelnefennqrcmcealqgimengsdrlqgaqqeqgfkrearnlpqg 134

QY 140 CGLRAPQRCDLQDVESGGR 157
||||| ||||||| |||||||
Db 135 cglrapqrccadvesggr 152

RESULT 10
AAU05035
ID AAU05035 standard; protein; 166 AA.
XX
AC AAU05035;
XX
DT 24-OCT-2001 (first entry)
XX
DE Modified anaphylactic peanut antigen Ara h 2.
XX
KW Anaphylactic antigen; Ara h 2; peanut; desensitisation; antigen;
KW allergy; immunotherapy; rush immunotherapy; anaphylaxis; asthma;
KW food allergen; milk allergen; shellfish allergen; latex; drug;
KW environmental allergen; grass pollen; ovalbumin; insect venom;
KW peanut oil; immunoglobulin E; IgE.
XX
OS Arachis sp.
OS Synthetic.
XX
PN WO200139799-A2.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US33125.
XX
PR 06-DEC-1999; 99US-0455294.
PR 23-JUN-2000; 2000US-0213765.
PR 27-SEP-2000; 2000US-0235797.
XX
PA (PANA-) PANACEA PHARM LLC.
XX
PI Caplan M;
XX
DR WPI: 2001-408258/43.
XX
PT Preventing allergic response against antigens, e.g. food and
PT environmental allergens such as peanut allergen or grass pollen, latex
PT or drug, comprises administering agent, e.g. a peptide, that blocks

PT antigen binding sites on offending IgE -
XX Disclosure; Fig 2; 76pp; English.
PS
XX
CC The sequence represents the amino acid sequence of modified anaphylactic
CC peanut antigen Ara h 2. The protein is used for active or passive
CC desensitisation of an individual to an antigen; for alleviating or
CC preventing allergic reactions and for decreasing the risk of allergic
CC reactions during immunotherapy or rush immunotherapy, anaphylaxis
CC and asthma. The antigen may be a food allergen (e.g. grass pollen
CC allergen), shellfish allergen, environmental allergen (e.g. grass pollen
CC or tree pollen), latex, drug, pollen, ovalbumin, an insect venom antigen
CC or predominantly linear epitopes. The protein is useful for protecting an
CC individual against subsequent inadvertent or intentional exposure to
CC antigen, e.g. receiving blocking agent before eating a chocolate bar
CC which may inadvertently contain peanut components or before eating foods
CC prepared using peanut oil. Administration of the blocking agents does not
CC result in cross-linking of anti-antigenic immunoglobulin E (IgE). After
CC exposure to the agent, the individual's antigen sensitivity is at least
CC temporarily reduced. Only those IgE molecules that bind the offending
CC antigen and contribute to the risk of an allergic response are blocked.
XX
SQ Sequence 166 AA;

Query Match 61.1%; Score 681; DB 22; Length 166;
Best Local Similarity 92.8%; Pred. No. 2.8e-64;
Matches 128; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 20 SARQQWELQGDRCQSLERANLRPCQHLQMQIKQDEDSYERDPYSPSPYDR 79
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 15 sarqqaqlgddrrcqsqqlaranlraceahlmqkqadedsyerapyspyspydr 74

QY 80 RGAGSSQHQRCCNELNEFENNQCRCMCEALQQIMENQSDRLQGRQQEQFKRELRLNPQQ 139
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 75 rgagssqhqrccnelnefennqrcmcealqgimengsdrlqgaqqeqgfkrearnlpqg 134

QY 140 CGLRAPQRCDLQDVESGGR 157
||||| ||||||| |||||||
Db 135 cglrapqrccadvesggr 152

RESULT 11
AAW23419
ID AAW23419 standard; Protein; 158 AA.
XX
AC AAW23419;
XX
DT 08-APR-1998 (first entry)
XX
DE Soybean albumin 3 protein.
XX
KW Soybean albumin 3; seed storage protein; SSP; transgenic seed;
KW transgenic soybean plant; animal feed production.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT Peptide 1..21 /note= "prepropeptide"
FT Protein 22..158 /note= "mature protein"
FT Misc-difference 22..81 /note= "small chain sequence"
FT Misc-difference 81..158 /note= "large chain sequence"
XX
PN WO9735023-A2.
XX
PD 25-SEP-1997.
XX
PF 19-MAR-1997; 97WO-US04409.
XX

XX FH Key Location/Qualifiers
FT Peptide 1..20
FT /note= "signal peptide"
FT Protein 21..155
FT /note= "mature protein"
FT Misc-difference 22..75
FT /note= "small chain sequence"
FT Misc-difference 76..155
FT /note= "large chain sequence"
XX

PN WO9735023-A2.
XX
XX 25-SEP-1997.
XX
XX 19-MAR-1997; 97WO-US04409.
XX
XX 20-MAR-1996; 96US-0618911.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Coughlan S, Hastings C, Hu D, Jung R;
PI WPI; 1997-480228/44.
XX N-PSDB; AAT86281.
DR
DR
XX

Increasing the nutritional quality of soybean - by genetically increasing the amount of lysine, methionine and/or cysteine in albumin proteins
Claim 25; Page 44-45; 62pp; English.

This sequence represents the soybean albumin 1 protein. DNA encoding this sequence represents a DNA molecule of the invention. The DNA molecules of the invention comprise a preselected DNA segment encoding a seed storage protein (SSP). (I) is used to produce transgenic seeds and plants, especially soybean plants that have increased levels of lysine and also increased levels of methionine and/or cysteine in albumin protein type 1 and/or 3. Increasing the levels of these amino acids increases the nutritional value of soybean produced. This is especially useful for producing animal feeds. The amount of lysine in the seed is increased by 5-10 %, the amount of methionine and cysteine is increased by 10-30 %

XX Sequence 155 AA;
SQ
Query Match 22.6%; Score 251.5; DB 18; Length 155;
Best Local Similarity 38.2%; Pred. No. 8.3e-19;
Matches 60; Conservative 30; Mismatches 54; Indels 13; Gaps 7;

QY 1 MAKLTILVALALFLAAHASARQQWELQGDRCQSLERANLRPCEQHLMQIKI---RDE 57
Db 1 mtkltilliallfi--ahtccaskwqhqgqescreqlkginlnpce-himekiqagrrge 57
QY 58 DSYERDPYSPQDPYSPSPYDRRGAGSSQ----HQERCCNELNEFENNQRCMCALQOIM 113
Db 58 dgsdedhillrtmp-grinyirkkegkeeeeghmqkccsenselk-spqcqckalqkim 115
QY 114 ENQSDRLQGRQEQQFKRELNLPPQCGLRAPQRCDL 150
Db 116 dnqseqlegk-ekkgmerelnmlnlaircrlgpmigcdl 151

RESULT 14
AAW23420
ID AAW23420 standard; Protein; 158 AA.
XX
XX AAW23420;
XX
DT 08-APR-1998 (first entry)
XX
DE Chimeric Soybean albumin 1/3 protein.
XX

KW Soybean albumin 1; seed storage protein; SSP; transgenic seed; transgenic soybean plant; animal feed production; chimeric protein.
KW
XX
OS Synthetic.
OS Glycine max.
XX
PN WO9735023-A2.
XX
PD 25-SEP-1997.
XX
PF 19-MAR-1997; 97WO-US04409.
XX
PR 20-MAR-1996; 96US-0618911.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Coughlan S, Hastings C, Hu D, Jung R;
XX
DR WPI; 1997-480228/44.
DR N-PSDB; AAT86283.

Increasing the nutritional quality of soybean - by genetically increasing the amount of lysine, methionine and/or cysteine in albumin proteins
Claim 25; Page 48-49; 62pp; English.

This sequence represents a chimeric soybean albumin protein. This sequence was created from the albumin 1 and albumin 3 protein sequences (see AAW23418 and AAW23419). The DNA encoding this sequence represents a DNA molecule of the invention. The DNA molecules of the invention comprise a preselected DNA segment encoding a seed storage protein (SSP). (I) is used to produce transgenic seeds and plants, especially soybean plants that have increased levels of lysine and also increased levels of methionine and/or cysteine in albumin protein type 1 and/or 3. Increasing the levels of these amino acids increases the nutritional value of soybean produced. This is especially useful for producing animal feeds. The amount of lysine in the seed is increased by 5-10 %, the amount of methionine and cysteine is increased by 10-30 %

XX Sequence 158 AA;
SQ
Query Match 21.5%; Score 239; DB 18; Length 158;
Best Local Similarity 35.4%; Pred. No. 1.8e-17;
Matches 56; Conservative 32; Mismatches 58; Indels 12; Gaps 6;

QY 1 MAKLTILVALALFLAAHASARQQWELQGDRCQSLERANLRPCEQHLMQIKI--RDED 58
Db 1 mtkftillililfci-ahtcsaskwqhqqd-scrkqlkgvnltpcekhimekiqgrgdd 58
QY 59 SYERDPYSPQDPYSPSPYDRRGAGSSQHQE-----RCCNELNEFENNQRCMCALQOI 112
Db 59 ddddddhillrtmrgirnyirkkegkdedeeegmqkccctemselk-spkcqckalqki 117
QY 113 MENQSDRLQGRQEQQFKRELNLPPQCGLRAPQRCDL 150
Db 118 menqseele-ekenkkmekelmlnlmtcmrfgpmigcdl 154

RESULT 15
AAW23586
ID AAW23586 standard; Protein; 158 AA.
XX
XX AAW23586;
XX
DT 30-SEP-1997 (first entry)
XX
DE Mabinlin MBLI from Capparis masaiikai.

XX MBL; sweetener; transgenic plant; recombinant protein; food;
KW beverage; animal feed; chewing gum; dental hygiene product;
KW pharmaceutical.

XX OS Capparidaceae
XX FH Key Location/Qualifiers
FT Peptide 1..39 /label= Signal
FT Protein 40..158 /label= MBLI
FT Region 40..70 /label= A_chain
FT Region 71..86 /note= "14 amino acids evidently lost when mature dimer is formed"
FT Region 87..158 /label= B_chain
XX PN WO9700945-A1.
XX PD 09-JAN-1997.
XX PF 21-JUN-1996; 96WO-US10669.
XX PR 23-JUN-1995; 95US-0000480.
XX PA (UYHA-) UNIV HAWAII.
XX PI Chen H, Hu Z, Sun SSM, Xiong L;
XX DR WPI; 1997-087372/08.
XX DR N-PSDB; AAT60775.
XX PT Mabinlin nucleotide sequences - for prodn. of transgenic plants or recombinant proteins, useful as sweeteners, partic. in foods
XX PS Disclosure; Fig 3; 38pp; English.
XX CC The present sequence represents the mabinlin MBLI. The encoding cDNA can modulate the expression of the gene encoding MBL in a cell or tissue.
CC Mabinlin comes from the seeds of the Capparis masakai plant, and is limited by its native availability, and so methods for producing recombinant mabinlin protein are being produced using chimeric genes such as MBLII (see AAT60776). The products provide recombinant materials for the production of the MBL protein in practical amounts, and for the production of transgenic plants containing inherently sweet edible parts by virtue of production of MBL in situ. The MBL protein can be used to sweeten foods, beverages, animal feeds, chewing gum, dental hygiene products and pharmaceuticals.
XX SQ Sequence 158 AA;

Query Match 14.6%; Score 162.5; DB 18; Length 158;
Best Local Similarity 29.8%; Pred. No. 2.3e-09;
Matches 50; Conservative 32; Mismatches 51; Indels 35; Gaps 9;
QY 1 MAKLTILV-ALALFLLAAHASA-RQWEL-----QGDRRCQSOLER-ANLRPCEQHLM 50
Db 1 maklililttialfvllanasirrttveldeednddenqplcrrfqgqhhracqyir 60
QY 51 QKIQRD--EDSYERDPYSPQDPYSPSPYDRRGAGSSQHQR-----CCNELNEFENNQ 102
Db 61 rraqrgglvdeledeveened-----edenqrgpalrlccnqlrqv--nk 105
QY 103 RCMCEALQIMENQ--SDRLQGRQEQQFKRELNLPPQCCGLRAPQRC 148
Db 106 pcvcvplrqaahqglyggiegprqvrqlfraarnlpnickipavgrc 153



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 13:03:11 ; Search time 8470.79 Seconds
(without alignments)
1328.220 Million cell updates/sec

Title: US-09-715-036-3

Perfect score: 682

Sequence: 1 gacacagacaaactgtaat.....ctgttgactgtagcggtga (682)

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vi:*

30: em_htgo_hum:*

31: em_htgo_inv:*

32: em_htgo_rod:*

33: em_htg_hum:*

34: em_htg_inv:*

35: em_htg_rod:*

36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|------------|--------------------|
| 1 | 682 | 100.0 | 682 | 6 | AX148740 | AX148740 Sequence |
| 2 | 682 | 100.0 | 1162 | 6 | AX148738 | AX148738 Sequence |
| 3 | 462 | 67.7 | 717 | 8 | ARQALLII | L77197 Arachis hyp |
| 4 | 459.4 | 67.4 | 474 | 6 | AX155333 | AX155333 Sequence |
| 5 | 157.4 | 23.1 | 627 | 8 | AF092846 | AF092846 Arachis h |
| 6 | 149.2 | 21.9 | 5787 | 12 | PEX2 | X03174 Cloning vec |
| 7 | 149.2 | 21.9 | 5787 | 12 | SYNPEX2V | L08868 PEX2 expres |
| 8 | 149.2 | 21.9 | 6728 | 12 | PUEX | X06291 Bacterial e |
| 9 | 146.8 | 21.5 | 1284 | 6 | I00684 | I00684 Sequence 12 |
| 10 | 146.8 | 21.5 | 3234 | 6 | A11623 | A11623 59 Val-IGF- |
| 11 | 146.8 | 21.5 | 3234 | 6 | A11624 | A11624 59 Val-IGF- |
| 12 | 146.8 | 21.5 | 3234 | 6 | E00609 | E00609 DNA encodin |
| 13 | 146.4 | 21.5 | 12294 | 12 | DMCASPLCZ | X81643 D.melanogas |
| 14 | 146 | 21.4 | 350 | 10 | S89402S1 | S89402 ferritin Lg |
| 15 | 146 | 21.4 | 1989 | 12 | CVSACKPN | X16973 Cloning vec |
| 16 | 146 | 21.4 | 3078 | 1 | ECLACZ | V00296 E. coli gen |
| 17 | 146 | 21.4 | 3150 | 6 | AX163849 | AX163849 Sequence |
| 18 | 146 | 21.4 | 3152 | 6 | A20699 | A20699 pMC LacZ DN |
| 19 | 146 | 21.4 | 3152 | 6 | I15648 | I15648 Sequence 3 |
| 20 | 146 | 21.4 | 3365 | 6 | I18302 | I18302 Sequence 1 |
| 21 | 146 | 21.4 | 3365 | 6 | I72358 | I72358 Sequence 1 |
| 22 | 146 | 21.4 | 3575 | 12 | SYNPRBGA | M81126 Artificial |
| 23 | 146 | 21.4 | 3624 | 12 | SYNBGPRA | M81125 Artificial |
| 24 | 146 | 21.4 | 3652 | 6 | AX114850 | AX114850 Sequence |
| 25 | 146 | 21.4 | 3652 | 6 | AX137076 | AX137076 Sequence |
| 26 | 146 | 21.4 | 4342 | 12 | SCU54830 | U54830 Synthetic E |
| 27 | 146 | 21.4 | 5019 | 6 | A11178 | A11178 Synthetic n |
| 28 | 146 | 21.4 | 5019 | 6 | E01007 | E01007 DNA sequenc |
| 29 | 146 | 21.4 | 5346 | 12 | SYNPUR288V | L09147 PUR288 clon |
| 30 | 146 | 21.4 | 5347 | 12 | SYNPUR289V | L09148 PUR289 clon |
| 31 | 146 | 21.4 | 5348 | 12 | SYNPUR291V | L09150 PUR291 clon |
| 32 | 146 | 21.4 | 5349 | 12 | SYNPUR292V | L09151 PUR292 clon |
| 33 | 146 | 21.4 | 5351 | 12 | SYNPUR278V | L09146 PUR278 clon |
| 34 | 146 | 21.4 | 5353 | 12 | SYNPUR290V | L09149 PUR290 clon |
| 35 | 146 | 21.4 | 5730 | 12 | SYNLACZD | M34299 Cloning vec |
| 36 | 146 | 21.4 | 5808 | 12 | SYNLACZL | M34307 Cloning vec |
| 37 | 146 | 21.4 | 5932 | 12 | U03993 | U03993 Cloning vec |
| 38 | 146 | 21.4 | 6096 | 12 | SYNLACZB | M34297 Cloning vec |
| 39 | 146 | 21.4 | 6107 | 12 | SCU54829 | U54829 Synthetic E |
| 40 | 146 | 21.4 | 6114 | 12 | CVU46489 | U46489 Cloning vec |
| 41 | 146 | 21.4 | 6119 | 12 | SCU54828 | U54828 Synthetic E |
| 42 | 146 | 21.4 | 6168 | 12 | CVU46490 | U46490 Cloning vec |
| 43 | 146 | 21.4 | 6252 | 12 | CVU46491 | U46491 Cloning vec |
| 44 | 146 | 21.4 | 6280 | 12 | XXU19930 | U19930 Cloning vec |
| 45 | 146 | 21.4 | 6476 | 12 | SYNLACZAB | M34296 Cloning vec |

ALIGNMENTS

RESULT 1

AX148740

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

AX148740

Sequence 3 from Patent WO0136621.

AX148740

AX148740.1 GI:14347294

peanut.

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;

Aeschynomeneae; Arachis.

1 (bases 1 to 682)

Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.

Down-regulation and silencing of allergen genes in transgenic

peanut seeds

Patent: WO 0136621-A 3 25-MAY-2001;

Alabama A & M University (US)

Location/Qualifiers

PAT

DNA

682 bp

GI:14347294

peanut.

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;

Aeschynomeneae; Arachis.

1 (bases 1 to 682)

Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.

Down-regulation and silencing of allergen genes in transgenic

peanut seeds

Patent: WO 0136621-A 3 25-MAY-2001;

Alabama A & M University (US)

Location/Qualifiers

08-JUN-2001

```

source      1. .682
            /organism="Arachis hypogaea"
            /db_xref="taxon:3818"
BASE COUNT  170 a  189 c  202 g  121 t
ORIGIN

Query Match      100.0%;   Score 682;   DB 6;   Length 682;
Best Local Similarity 100.0%;   Pred. NO. 1.8e-171;
Matches 682;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY  1  gacacagaccaactgtaagttagcgaccggcgctcagctggaattcgcggccgccaat 60
      |||
Db  1  GACACAGACCAACTGTAATGTTAGCAGACCGCGCTCAGCTGGAATTCGCGGCCGCCAAT 60

QY  61  ggccaagctcaccatactagtagccctcgccctttctcctcgtcgctgccacgcatctgc 120
      |||
Db  61  GGCCAAGCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTGC 120

QY  121  gaggcagcagtggaactcaaggagacagaagatgccagagccagctcgagaggcgaa 180
      |||
Db  121  GAGGCAGCAGTGGGAACCTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAA 180

QY  181  cctgaggccctgcgagcaacatctcatgcagaagatccaacgtgacgaggttcatatga 240
      |||
Db  181  CCTGAGGCCCTGGGAGCAACATCTCATGCAGAAGATCCACCGTGACGAGGATTCATATGA 240

QY  241  acgggacccgtacagccctagtcaggatccgtacagccctagtcctcatatgatcggagagg 300
      |||
Db  241  ACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGG 300

QY  301  cgctggatcctctcagcaccgaagagaggtgttgcaatgagctgaacagagtttgagaacaa 360
      |||
Db  301  CGCTGGATCCTCTCAGCACCAAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACAA 360

QY  361  ccaaaggtgcatgtgagggcattgcaacagatcatggagaaaccagagcgataggttgca 420
      |||
Db  361  CCAAAGGTGCATGTGCGAGGGCATTGCAACAGATCATGGAGAACCAAGAGCGATAGGTTGCA 420

QY  421  ggggaggcaacaggagcaacagtccaagaggagctcaggaaacttgccctcaacagtcgag 480
      |||
Db  421  GGGGAGGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCTCAACAGTCGGG 480

QY  481  ccttagggcaccacagcgttgagacttgagctcgaaagtggcgagggcgagccgcaaat 540
      |||
Db  481  CCTTAGGGCACCAACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGAGGCGCGCGCAAT 540

QY  541  tcgcccgatactgacggggtccaggagtcgtcgccaccaatcccccatatggaaaccgtcg 600
      |||
Db  541  TCCGCCGATACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCCCATATGGAAACCGTCG 600

QY  601  atattcagccatgtgccttcttcggcgtgcagcagatggcgatggctgggttcccatcagt 660
      |||
Db  601  ATATTAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTGTTCATCAGT 660

QY  661  tgctgttgactgtagcggctga 682
      |||
Db  661  TGCTGTGACTGTAGCGGCTGA 682

```

RESULT 2
AX148738
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AX148738
Sequence 1 from Patent WO0136621.
AX148738
AX148738.1 GI:14347292
peanut.
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.


```
misc_feature      1195. .4140
                  /note="lacZ fragment"
misc_feature      4141. .4176
                  /note="cloning linker"
misc_feature      4142. .4145
                  /note="AATT deleted in pUEX1"
misc_feature      4145. .4146
                  /note="AATT inserted in pUEX3"
misc_feature      4179. .4189
                  /note="stop codon region"
misc_feature      4214. .4316
                  /note="transcription terminator 1 from phage fd"
misc_feature      4323. .4425
                  /note="transcription terminator 2 from phage"

BASE COUNT      1629 a 1735 c 1727 g 1637 t
ORIGIN

Query Match      21.9%; Score 149.2; DB 12; Length 6728;
Best Local Similarity 95.1%; Pred. No. 2.7e-29;
Matches 154; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 521 ggcggcaggccgcgaattccgcgatactgacgggtccaggagtcgtgccaccaa 580
    |||||
Db 4162 GTCGACGGATCCCGGGGAATCCGCGGATACTGACGGGCTCCAGGAGTCGTCGCCACCAA 4103

QY 581 tccccatatgaaacccgtcgatattcagccatgtgccttctccgcgtgcagagatggc 640
    |||||
Db 4102 TCCCCATATGAAACCGTCGATATTAGCCCATGTGCCCTCTTCCGCGTGCAGCAGATGGC 4043

QY 641 gatggctgttccatcagttcgtgtgactgtagcggctga 682
    |||||
Db 4042 GATGGCTGTTCCATCAGTTGCTGTGACTGTAGCGGCTGA 4001

RESULT 9
LOCUS      I00684      1284 bp ss-DNA      PAT      21-MAY-1993
DEFINITION      Sequence 12 from Patent US 4745179.
ACCESSION      I00684
VERSION      I00684.1 GI:269096
KEYWORDS      .
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 1284)
AUTHORS      Ueda,I., Niwa,M., Saito,Y., Sato,S., Ono,H. and Kitaguchi,T.
TITLE      .sup.59 Valine insulin-like growth factor I and process for
          production thereof
JOURNAL      Patent: US 4745179-A 12 17-MAY-1988;
          Fujisawa Pharmaceutical Co., Ltd.;
          Osaka, ;
          JP;

FEATURES
    source      Location/Qualifiers
                1..1284
                /organism="unknown"

BASE COUNT      284 a 360 c 372 g 268 t
ORIGIN

Query Match      21.5%; Score 146.8; DB 6; Length 1284;
Best Local Similarity 98.7%; Pred. No. 1.4e-28;
Matches 148; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 533 ccgcgaattccgcgatactgacgggctccaggagtcgtgccaccaatccccatatgga 592
    ||
Db 1075 CCATGAATTCGCCGATACTGACGGGCTCCAGGAGTCGTCGCCACCAATCCCATATGGA 1016

QY 593 aaccgtcagattcagccatgtgccttctccgcgtgcagcagatggcgatggctggtt 652
    |||||
Db 1015 AACCGTCGATATTCAGCCATGTGCTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTT 956

QY 653 ccatcagttgctgtgactgtagcggctga 682
```

```
|||||
Db 955 CCATCAGTTGCTGTTGACTGTAGCGGCTGA 926

RESULT 10
LOCUS      A11623/c      3234 bp      DNA      PAT      17-NOV-1993
DEFINITION      59 Val-IGF-I fused with beta galactosidase.
ACCESSION      A11623
VERSION      A11623.1 GI:491150
KEYWORDS      .
SOURCE      synthetic construct.
ORGANISM      synthetic construct
          artificial sequence.
REFERENCE      1 (bases 1 to 3234)
AUTHORS      Ueda,I., Niwa,M., Saito,Y., Sato,S., Ono,H. and Kitaguchi,T.
TITLE      59 Valine insulin-like growth factor I and process for production
          thereof
JOURNAL      Patent: EP 0158892-A 119 23-OCT-1985;
          FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
    source      Location/Qualifiers
                1..3234
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                1..3234
                /gene="59val-IGF-I fused with beta galactosidase"
                1..3234
                /gene="59val-IGF-I fused with beta galactosidase"
                /note="protein sequence is in conflict with the conceptual
          translation"
                /codon_start=1
                /transl_table=11
                /protein_id="CAA00972.1"
                /db_xref="GI:4526654"
                /translation="TMTDSLAVVLQRRDWENPGVTQLNRLAAHPPFASWRNSEART
          DRPSQLRLSNGEWFAPFAPEAVPESWLECDLPEADTVVVPNSWQHGYDAPITYN
          VTYPTIVNPPFVPTENTPGYSLTFNVDESWLQEGQTRIIDGVNSAFHLWCNGRWVG
          YQDSRLPSEFDLSAFLRAGENRLAVMLRWSDSYLEDQDMWRMSGIFRDVSLHKP
          TTIQISDFHVAIRFNDDFSRAVLEAEVQCMGELRDYLRVTVSLWQGETVASGTAPFGG
          EIIDERGAYADRVTLRLNVENPKLWSAEIPNLRAVVELHTADGTLIEACDVGFRE
          VRIENGLLLNGKPLLRGVNRHEHPLHGQVMDQTMVDILLMKNNFNNAVRCSHY
          NHPMLWTLCDRYGLYVDEANIETHGMVPMNRLTDDPRWLPAMSERVTRMVQDRNH
          PSVITISLGNESGHGANHDALYRWIKSVDPSPVQYEGGADTTATDIICPMYARVDE
          DQPFPAVPKWSIKKWLSPGETRPLILCEYAHAMGNSLGGFAKYWQAFRQYPRLOGGF
          VWDVVDOSLIKYDENGPNWSAYGGDFGDPNDROFCMNGLVFADRTPHPALTEAKHQQ
          QFFQRLSGQTIETVSEYLRHSDNELLHWMVALDGKPLASGEVPLDVAPOGKQLIEL
          PELPOPSAQQLWLTVRVVOPNATAWSEAGHSIAWQWRLAENLSVTLPAASHAIPHL
          TTSEMDFCIELGNKRWFNRQSGFLSQMWIGKKQLTLPLRDQFTRAPLDNDIGVSEA
          TRIDPNAWVERWKAAGHYQAEALLQCTADTLADAVLITTAHWQHOGKTLFISRKTY
          RIDGSCQMAITVDVEVADTPHPARIGLNCQLAQAVERNWLGLPQENYPDRLTAAC
          FDRWDLPLSDMYTPYVFPSENGLRCGTRELNYPGHQWRGDFQFNISYSQQQLMETSH
          RHLHAEEGTWLNIDGFHMGIGGDDSSWSPSVAEFMGPETLCGAELVDALQFVCGDRG
          FYFNKPTGYGSSRRRAPQTGIVDECCFRSCDLRLLEVYCAPLKPAKSA"

BASE COUNT      705 a 886 c 929 g 714 t
ORIGIN

Query Match      21.5%; Score 146.8; DB 6; Length 3234;
Best Local Similarity 98.7%; Pred. No. 1.3e-28;
Matches 148; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 533 ccgcgaattccgcgatactgacgggctccaggagtcgtgccaccaatccccatatgga 592
    ||
Db 3025 CCATGAATTCGCCGATACTGACGGGCTCCAGGAGTCGTCGCCACCAATCCCATATGGA 2966

QY 593 aaccgtcagattcagccatgtgccttctccgcgtgcagcagatggcgatggctggtt 652
    |||||
Db 2965 AACCGTCGATATTCAGCCATGTGCTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTT 2906

QY 653 ccatcagttgctgtgactgtagcggctga 682
    |||||
Db 2905 CCATCAGTTGCTGTTGACTGTAGCGGCTGA 2876
```

RESULT 11
A11624
LOCUS
DEFINITION 59 Val-IGF-I fused with beta galactosidase. PAT 17-NOV-1993
ACCESSION A11624
VERSION A11624.1 GI:489368
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3234)
AUTHORS Ueda,I., Niwa,M., Saito,Y., Sato,S., Ono,H. and Kitaguchi,T.
TITLE 59 Valine insulin-like growth factor I and process for production thereof
JOURNAL Patent: EP 0158892-A 120 23-OCT-1985;
FUJISAWA PHARMACEUTICAL CO., LTD
FEATURES
source Location/Qualifiers
1. .3234
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 714 a 929 c 886 g 705 t
ORIGIN
Query Match 21.5%; Score 146.8; DB 6; Length 3234;
Best Local Similarity 98.7%; Pred. No. 1.3e-28;
Matches 148; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 533 ccgcgaattccgccgatactgacgggctccaggagtcgtcgccaccaatccccatatgga 592
|| |||||
Db 210 CCATGAATTCGCCGCGATACTGACGGGCTCCAGGAGTCGTCCGCCACCAATCCCCATATGGA 269
QY 593 aaccgtcgatatccagccatgtgccttcttcgcgctgcagcagatggcgatggctggttt 652
|| |||||
Db 270 AACCGTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTT 329
QY 653 ccatcagttgctgttgactgtagcggctga 682
|| |||||
Db 330 CCATCAGTTGCTGTTGACTGTAGCGGCTGA 359
RESULT 12
E00609/c
LOCUS
DEFINITION E00609 DNA encoding (beta-gal)-(59 Val-IGFI) fused protein. PAT 29-SEP-1997
ACCESSION E00609
VERSION E00609.1 GI:2168888
KEYWORDS JP 1986001397-A/3.
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 3234)
AUTHORS Ueda,I., Niwa,M., Saito,Y., Sato,S., Ono,H. and Kitaguchi,T.
TITLE GROWTH FACTOR I-LIKE 59VALINEINSULIN AND ITS PREPARATION
JOURNAL Patent: JP 1986001397-A 3 07-JAN-1986;
FUJISAWA PHARMACEUT CO LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986001397-A/3
PD 07-JAN-1986
PF 01-APR-1985 JP 1985069630
PR 02-APR-1984 GB 84 8408473, 01-JUN-1984 GB 84 8413989, PR
25-SEP-1984 GB 84 8424157
PI UEDA IKUO, NIWA MINEO, SAITO YOSHIMASA, SATO SUSUMU, PI ONO
HIROKI,
PI KITAGUCHI TADASHI
PC C12P21/02,A61K35/74,A61K37/24,A61K37/26,C07H21/04,C12N1/00, PC
C12N15/00,
PC (C12P21/02,C12R1:19),(C12N1/00,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;

CC anti-sense: No; Location/Qualifiers
FH Key
FH CDS <1..>3234
FT /product="(beta-gal)-(59 Val-IGFI)fused FT
FT protein'
FT misc_feature 1..3021
FT /note='beta-gal coding region' FT
FT misc_feature 3025..3234
FT /note='59 Val-IGFI coding region'.
FEATURES
source Location/Qualifiers
1. .3234
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 705 a 889 c 926 g 714 t
ORIGIN
Query Match 21.5%; Score 146.8; DB 6; Length 3234;
Best Local Similarity 98.7%; Pred. No. 1.3e-28;
Matches 148; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 533 ccgcgaattccgccgatactgacgggctccaggagtcgtcgccaccaatccccatatgga 592
|| |||||
Db 3025 CCATGAATTCGCCGCGATACTGACGGGCTCCAGGAGTCGTCCGCCACCAATCCCATATGGA 2966
QY 593 aaccgtcgatatccagccatgtgccttcttcgcgctgcagcagatggcgatggctggttt 652
|| |||||
Db 2965 AACCGTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTT 2906
QY 653 ccatcagttgctgttgactgtagcggctga 682
|| |||||
Db 2905 CCATCAGTTGCTGTTGACTGTAGCGGCTGA 2876
RESULT 13
DMCASPLCZ/c
LOCUS DMCASPLCZ 12294 bp DNA SYN 29-APR-2000
DEFINITION D.melanogaster P element CasPer-hs43-lacZ gene transformation vector.
ACCESSION X81643
VERSION X81643.1 GI:551429
KEYWORDS Carnegie-4 vector; P element; transformation vector; white gene.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 12294)
AUTHORS Pirrotta,V. and Zeng,C.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 12294)
AUTHORS Pirrotta,V.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-1994) V. Pirrotta, Dept of Zoology, University of Geneva, 30 Quai E. Ansermet, 1211 Geneva, SWITZERLAND
COMMENT Related sequences: M28731, M30841, X02974.
FEATURES
source Location/Qualifiers
1. .12294
/organism="synthetic construct"
/db_xref="taxon:32630"
/focus
source 1. .12294
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
source 1. .12294
/organism="Escherichia coli"
/db_xref="taxon:562"
misc_feature 1. .104
/note="polylinker"
promoter 111..245
/gene="hs43"
source order(111..245,290..417,4625..5366,7396..8208,8220..12294)
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

entry [NCBI gibbsq 89402] from the original journal article.
This sequence comes from Figure 5.

FEATURES
source
Location/Qualifiers
1. .350
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 85 a 99 c 88 g 77 t 1 others
ORIGIN

Query Match 21.4%; Score 146; DB 10; Length 350;
Best Local Similarity 100.0%; Pred. No. 2.6e-28;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggctccaggagtcgctgccaccacaaatccccatatggaaacc 596
|||||
Db 265 GAATTCCGCCGATACTGACGGGCTCCAGGAGTGGTCGCCACCAATCCCATATGGAACC 206

QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 205 GTCGATATTGAGCCATGTGCCTTCTTCCGCGTCGAGCAGATGGCGATGGCTGGTTTCCAT 146

QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 145 CAGTTGCTGTTGACTGTAGCGGCTGA 120

RESULT 15
CVSACKPN/C
LOCUS CVSACKPN 1989 bp DNA SYN 15-NOV-1990
DEFINITION Cloning vector lambda-gtl1 DNA sequence of SacI-KpnI region.
ACCESSION X16973
VERSION X16973.1 GI:58246
KEYWORDS cloning vector.
SOURCE Cloning vector lambda gtl1.
ORGANISM Cloning vector lambda gtl1
artificial sequence; vectors.
REFERENCE 1 (bases 1 to 1989)
AUTHORS Slatko,B.E.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1989) Slatko B. E., New England Biolabs, Inc., 32
Tozer Road Beverly MA 01915, USA
REFERENCE 2 (bases 1 to 1989)
AUTHORS Moran,L.S., Maina,C.V., Poole,C.B. and Slatko,B.E.
TITLE Nucleotide sequence of the phage lambda gtl1 SacI-KpnI lacz region
JOURNAL Gene 93 (1), 163-164 (1990)
MEDLINE 91033010
COMMENT See <X04125> for lambda gtl1 lac operon and phage junction. Data
kindly reviewed (15-DEC-1989) by Slatko B.

FEATURES
source
Location/Qualifiers
1. .1989
/organism="Cloning vector lambda gtl1"
/db_xref="taxon:31786"
/clone="pBS:LUC8"
/clone_lib="rest frags SacI-KpnI double digest"
misc_feature 1. .6
/note="SacI restriction site"
misc_feature 1069. .1074
/note="EcoRI restriction site"
misc_feature 1984. .1989
/note="KpnI restriction site"
BASE COUNT 471 a 597 c 518 g 403 t
ORIGIN

Query Match 21.4%; Score 146; DB 12; Length 1989;
Best Local Similarity 100.0%; Pred. No. 2.2e-28;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggctccaggagtcgctgccaccacaaatccccatatggaaacc 596
|||||
Db 1074 GAATTCGCCGATACTGACGGGCTCCAGGAGTGGTCGCCACCAATCCCATATGGAACC 1015

QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 1014 GTCGATATTGAGCCATGTGCCTTCTTCCGCGTCGAGCAGATGGCGATGGCTGGTTTCCAT 955

QY 657 cagttgctgttgactgttagcggctga 682
|||||
Db 954 CAGTTGCTGTTGACTGTAGCGGCTGA 929

Search completed: December 27, 2001, 13:03:21
Job time: 13867 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 13:14:47 ; Search time 486.97 Seconds
(without alignments)
1200.681 Million cell updates/sec

Title: US-09-715-036-3

Perfect score: 682

Sequence: 1 gacacagacaaactggtaat.....ctgttgactgtagcgctga 682

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101:*

- 1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 682 | 100.0 | 682 | 22 | AAF90337 |
| 2 | 682 | 100.0 | 1162 | 22 | AAF90336 |
| 3 | 462 | 67.7 | 717 | 18 | AAT76614 |
| 4 | 462 | 67.7 | 717 | 18 | AAT76615 |
| 5 | 461 | 67.6 | 474 | 20 | AAS06383 |
| 6 | 461 | 67.6 | 474 | 22 | AAS08538 |
| 7 | 146 | 21.4 | 3150 | 22 | AAH45081 |
| 8 | 146 | 21.4 | 3365 | 14 | AAQ41287 |
| 9 | 146 | 21.4 | 3652 | 22 | AAD04924 |
| 10 | 146 | 21.4 | 3652 | 22 | AAF85070 |
| 11 | 146 | 21.4 | 3756 | 20 | AAH87978 |
| | | | | | Peanut allergen Ar |
| | | | | | Peanut allergen Ar |
| | | | | | Peanut allergen Ar |
| | | | | | Peanut allergen Ar |
| | | | | | Peanut allergen, A |
| | | | | | DNA encoding anaph |
| | | | | | LacZ gene. Escher |
| | | | | | Ubiquitin fusion p |
| | | | | | E. coli beta-galac |
| | | | | | Nucleotide sequenc |
| | | | | | HCMV UL45 promoter |

| | | | | | | |
|------|-----|------|------|----|----------|---------------------|
| C 12 | 146 | 21.4 | 4279 | 20 | AAH06800 | Multiple cloning s |
| C 13 | 146 | 21.4 | 4487 | 21 | AAA09085 | AdRSV-beta-galacto |
| C 14 | 146 | 21.4 | 5201 | 21 | AAA09087 | AdPB-beta-galactos |
| C 15 | 146 | 21.4 | 5346 | 21 | AAA09089 | AdPSA-beta-galacto |
| C 16 | 146 | 21.4 | 5835 | 21 | AAA09091 | AdMTV-beta-galact |
| C 17 | 146 | 21.4 | 6050 | 22 | AAD10237 | Commercial plasmid |
| C 18 | 146 | 21.4 | 6560 | 20 | AAH26118 | DNA sequence of FP |
| C 19 | 146 | 21.4 | 6561 | 20 | AAH26117 | DNA sequence of FP |
| C 20 | 146 | 21.4 | 6700 | 22 | AAH74523 | Nucleotide sequenc |
| C 21 | 146 | 21.4 | 6796 | 13 | AAQ27433 | HVTA antigen (ga) |
| C 22 | 146 | 21.4 | 6995 | 16 | AAQ85466 | Ras oncogene/LacZ |
| C 23 | 146 | 21.4 | 7123 | 20 | AAH86929 | Complete sequence |
| C 24 | 146 | 21.4 | 7174 | 10 | AAH91062 | Sequence of plasmid |
| C 25 | 146 | 21.4 | 7175 | 22 | AAD04927 | Vector PRK57SA-bet |
| C 26 | 146 | 21.4 | 7252 | 20 | AAH86928 | Complete sequence |
| C 27 | 146 | 21.4 | 7319 | 19 | AAV60246 | Vector VP1380 cont |
| C 28 | 146 | 21.4 | 7573 | 22 | AAD04930 | Recombination vect |
| C 29 | 146 | 21.4 | 7573 | 22 | AAD04944 | PRK73-inv vector c |
| C 30 | 146 | 21.4 | 8062 | 22 | AAD04946 | Plasmid PRK76 used |
| C 31 | 146 | 21.4 | 8136 | 18 | AAH63236 | Plasmid PCR(TM)31a |
| C 32 | 146 | 21.4 | 8153 | 22 | AAD04945 | Plasmid PRK74 used |
| C 33 | 146 | 21.4 | 8518 | 22 | AAH74524 | Nucleotide sequenc |
| C 34 | 146 | 21.4 | 8574 | 20 | AAH77158 | Nucleic acid seque |
| C 35 | 146 | 21.4 | 8657 | 19 | AAV33614 | Plasmid VLSNOSIB D |
| C 36 | 146 | 21.4 | 8710 | 16 | AAQ89650 | PSC11 FIPV E1. Sy |
| C 37 | 146 | 21.4 | 9020 | 16 | AAT07652 | PSC11 FIPV N. Syn |
| C 38 | 146 | 21.4 | 9093 | 19 | AAV33623 | Plasmid pVLMB DNA |
| C 39 | 146 | 21.4 | 9093 | 20 | AAH77157 | Nucleic acid seque |
| C 40 | 146 | 21.4 | 9138 | 22 | AAH83047 | ETAV expression v |
| C 41 | 146 | 21.4 | 9146 | 20 | AAH77156 | Nucleic acid seque |
| C 42 | 146 | 21.4 | 9641 | 20 | AAH15626 | pGEM Age Pac-beta- |
| C 43 | 146 | 21.4 | 9729 | 22 | AAH83093 | Nucleotide sequenc |
| C 44 | 146 | 21.4 | 9756 | 20 | AAH11457 | Seq ID NO: 2 of WO |
| C 45 | 146 | 21.4 | 9756 | 20 | AAH61062 | Retroviral vector |

ALIGNMENTS

RESULT 1

AAF90337

ID AAF90337 standard; DNA; 682 BP.

XX AAF90337;

AC AAF90337;

XX 23-JUL-2001 (first entry)

DT Peanut allergen Ara h2 gene.

XX Peanut; allergen; Ara h2; Ara h6; Ara h7; transgenic plant;

DE allergy; ds.

XX Arachis hypogea.

OS Arachis hypogea.

XX Key

FT CDS

FT sig_peptide

FT mat_peptide

FT misc_feature

FT WO200136621-A2.

FT 25-MAY-2001.

XX 20-NOV-2000; 2000WO-US31657.

XX 19-NOV-1999; 99US-0167255.

XX

Answer

X

|||||
Db 301 gtgcatgtgcgagcattgcaacagatcatggagaaccagagcgataggttgcaggaggag 360
QY 427 gcaacagaggcaacagttaagaggagctcaggaactgcctcaacagtgcggccttag 486
Db 361 gcaacagaggcaacagttaagaggagctcaggaactgcctcaacagtgcggccttag 420
QY 487 ggcaccacagcgttgcgacttgacgtcgaaagtggcggcag 528
Db 421 ggcaccacagcgttgcgacttgacgtcgaaagtggcggcag 462

RESULT 4
AAT76615
ID AAT76615 standard; cDNA to mRNA; 717 BP.
XX
AC AAT76615;
DT 29-DEC-1997 (first entry)
XX Peanut allergen Ara hII cDNA clone.
DB
XX peanut-seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy;
KW monoclonal antibody; ELISA; analysis; Ara hII; ds.
XX
OS Arachis hypogaea strain Florunner.

Key Location/Qualifiers
CDS 2..475
FT polyA_signal 562..567
FT /*tag= b
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US15222.
XX
PR 04-MAR-1996; 96US-0610424.
PR 29-DEC-1995; 95US-0009455.
XX
PA (UYAR-) UNIV ARKANSAS.

ant ☆

PI Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
XX
DR WPI; 1997-363453/33.
DR P-PSDB; AAW24153.
XX
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in
PT two-site monoclonal antibody based ELISA

PS Claim 31; Page 196; 354pp; English.
XX
CC This cDNA clone codes for the major peanut allergen Ara hII
CC (AAW22153), which contains multiple IgE-binding epitopes. It was
CC amplified from a peanut seed cDNA library using a primer (see
CC AAT58683) based on an isolated Ara hII peptide (see AAW24151). The
CC sequence shows significant homology with the conglutin family of
CC seed storage proteins of other legumes. The gene is capable of
CC producing a protein product in prokaryotic cells that is recognised
CC by serum IgE from a large proportion of individuals with peanut
CC hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used
CC to raise monoclonal antibodies which are used in a specific two-
CC site MAb ELISA for the detection of Ara hI or Ara hII (claimed).
CC IgE-binding Ara hII antigen epitopes (see AAW24188- 93) may be used
CC in vaccines to protect against allergic reactions to peanut
CC allergens, e.g. anaphylactic shock.

XX Sequence 717 BP; 217 A; 152 C; 184 G; 164 T; 0 other;

Query Match 67.7%; Score 462; DB 18; Length 717;
Best Local Similarity 100.0%; Pred. No. 3.4e-126;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 gctcaccatactagtagccctcgcccttttctcctcgctgcccacgcattcgcgaggca 126
Db 1 gctcaccatactagtagccctcgcccttttctcctcgctgcccacgcattcgcgaggca 60
QY 127 gcagtgggaaactccaaggagacagaagatgccagagccagctcgagaggcgaaacctgag 186
Db 61 gcagtgggaaactccaaggagacagaagatgccagagccagctcgagaggcgaaacctgag 120
QY 187 gccctgcgagcaacatctcatgcagaagatccaaacgtgacgaggttcatatgaacggga 246
Db 121 gccctgcgagcaacatctcatgcagaagatccaaacgtgacgaggttcatatgaacggga 180
QY 247 cccgtacagcccttagtcaggatccgtacagcccttagtcacatgacgcgagagcgctgg 306
Db 181 cccgtacagcccttagtcaggatccgtacagcccttagtcacatgacgcgagagcgctgg 240
QY 307 atcctctcagcaccagagaggtgtgtgcaatgagctgaacgagtttgagaacaaacaaag 366
Db 241 atcctctcagcaccagagaggtgtgtgcaatgagctgaacgagtttgagaacaaacaaag 300
QY 367 gtgcatgtgcgaggttcgaacagatcatggagaacagagcgataggttgcagggggag 426
Db 301 gtgcatgtgcgaggttcgaacagatcatggagaacagagcgataggttgcagggggag 360
QY 427 gcaacagaggcaacagttcaagaggggagctcaggaacttgcctcaacagtcggccttag 486
Db 361 gcaacagaggcaacagttcaagaggggagctcaggaacttgcctcaacagtcggccttag 420
QY 487 ggcaccacagcgttgcgacttggaacttcgaaagtggcggcag 528
Db 421 ggcaccacagcgttgcgacttggaacttcgaaagtggcggcag 462

RESULT 5
AAZ06383
ID AAZ06383 standard; DNA; 474 BP.
XX
AC AAZ06383;
XX 09-NOV-1999 (first entry)
DT
XX Peanut allergen, Ara h 2.
DE
XX allergy; immune response; transgenic; allergen; epitope;
KW immunoglobulin E; Ig E; binding site; peanut; ds.
KW
XX
OS Arachis hypogaea.
XX
PN WO9938978-A1.
XX
PD 05-AUG-1999.
XX
PF 29-JAN-1999; 99WO-US02031.
XX
PR 27-AUG-1998; 98US-0141220.
PR 31-JAN-1998; 98US-0073283.
PR 13-FEB-1998; 98US-0074590.
PR 13-FEB-1998; 98US-0074624.
PR 13-FEB-1998; 98US-0074633.
XX
PA (SOSI/) SOSIN H.
PA (UYAR-) UNIV ARKANSAS.
PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

ant ☆

PI Bannon GA, Burks AW, Sampson HA, Sosin H;
XX
DR WPI; 1999-479189/40.
DR P-PSDB; AAY15245.
XX

PT Modified allergen with reduced IgE binding, useful for treating e.g. allergies

XX

PS

XX

XX

Disclosure; Page 37; 46pp; English.

CC This is the nucleotide sequence of the Ara h 2 protein from Arachis hypogaea. The Ara h 2 protein has 10 IgE (Immunoglobulin E) binding epitopes, three of which are immunodominant (AA15272, AA15275, and AA15276).

CC By modifying the IgE binding sites the ability of the allergen to provoke an immune response is downregulated. The epitopes of the IgE binding sites can therefore be modified in genetically engineered plants and animals to elicit less of an allergic response.

XX

SQ Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 other;

Query Match 67.6%; Score 461; DB 20; Length 474;
Best Local Similarity 100.0%; Pred. No. 5.7e-126;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ctcaccatactagtagccctcgcccttttctctcgtgctgccacgcatctcgaggcag 127
|||||

Db 1 ctcaccatactagtagccctcgcccttttctctcgtgctgccacgcatctcgaggcag 60

QY 128 cagtgggaactccaaggagacagaagatgccagagccagctcgagaggcgaaacctgagg 187
|||||

Db 61 cagtgggaactccaaggagacagaagatgccagagccagctcgagaggcgaaacctgagg 120

QY 188 ccctgcgagcaacatctcatgcagaagatccaacgtgacgaggattcatatgaacgggac 247
|||||

Db 121 ccctgcgagcaacatctcatgcagaagatccaacgtgacgaggattcatatgaacgggac 180

QY 248 ccgtacagccctagtcaggatccgtacagccctagtcctcatatgatcggagagcgctgga 307
|||||

Db 181 ccgtacagccctagtcaggatccgtacagccctagtcctcatatgatcggagagcgctgga 240

QY 308 tcctctcagcaccacagagaggtgttgcaatgagctgaacgagtttgagaacaaccaaagg 367
|||||

Db 241 tcctctcagcaccacagagaggtgttgcaatgagctgaacgagtttgagaacaaccaaagg 300

QY 368 tgcattgtgcgaggcattgcaacagatcatcgagaaaccagagcgataggttcaggggagg 427
|||||

Db 301 tgcattgtgcgaggcattgcaacagatcatcgagaaaccagagcgataggttcaggggagg 360

QY 428 caacaggagcaacagttcaaggaggagctcaggaaacttgcctcaacagtgcggccttagg 487
|||||

Db 361 caacaggagcaacagttcaaggaggagctcaggaaacttgcctcaacagtgcggccttagg 420

QY 488 gcaccacagcgttgagacttgacgctcgaaagtggcggcag 528
|||||

Db 421 gcaccacagcgttgagacttgacgctcgaaagtggcggcag 461

RESULT 6

AAS08538

ID AAS08538 standard; cDNA; 474 BP.

XX

AC AAS08538;

XX

XX 23-OCT-2001 (first entry)

DT

XX DNA encoding anaphylactic antigen Ara h 2.

DE

XX Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic; allergy; mast cell; basophil; mouse; ss.

KW

KW

XX

OS Mus sp.

XX

XX

FH Key

CD Location/Qualifiers

1..474

/*tag= a

/product= "Peptide antigen Ara h 2"

FT

XX WO200140264-A2.

PN

XX

PD 07-JUN-2001.

XX

PF 06-DEC-2000; 2000WO-US33124.

XX

PR 06-DEC-1999; 99US-0455294.

PR 23-JUN-2000; 2000US-0213765.

PR 27-SEP-2000; 2000US-0235797.

XX

PA (PANA-) PANACEA PHARM LLC.

PA (UYAR-) UNIV ARKANSAS.

PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.

XX

PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;

XX

DR WPI; 2001-381378/40.

DR P-PSDB; AAU04707.

XX

PT Antigenic fragments useful for reducing anaphylactic risk and reducing the severity and/or number of allergic symptoms in individuals sensitive to antigens, have reduced ability to bind Immunoglobulin E -

PT

XX

PS Disclosure; Fig 10; 100pp; English.

XX

CC The sequence represents the coding sequence of anaphylactic antigen Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design antigenic peptides having a reduced ability to bind immunoglobulin E (IgE) as compared with the intact (A), or having a sequence substantially identical to a portion of sequence of an antigen that includes at least one IgE binding site, where at least one IgE binding site of the peptide is altered. The antigenic peptides are used in a composition which is useful for reducing risk or severity of allergic reaction to an antigen. This is done by identifying an individual at risk of allergic reaction to an antigen by identifying prior display of allergic symptoms when exposed to the antigen, or a familial relationship with an individual who previously displayed allergic symptoms when exposed to the antigen. Following this an antigen-specific IgE present on one or more mast cells or basophils in the individual's serum is identified. The individual is then contacted with a peptide corresponding to a portion of the antigen, which is selected, formulated, and delivered so that binding of the peptide to antigen-specific IgE is reduced as compared with IgE binding of intact antigen. The composition is also useful for treating and preventing allergic reactions.

CC

XX

SQ Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 other;

Query Match 67.6%; Score 461; DB 22; Length 474;
Best Local Similarity 100.0%; Pred. No. 5.7e-126;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 ctcaccatactagtagccctcgcccttttctctcgtgctgccacgcatctcgaggcag 127
|||||

Db 1 ctcaccatactagtagccctcgcccttttctctcgtgctgccacgcatctcgaggcag 60

QY 128 cagtgggaactccaaggagacagaagatgccagagccagctcgagaggcgaaacctgagg 187
|||||

Db 61 cagtgggaactccaaggagacagaagatgccagagccagctcgagaggcgaaacctgagg 120

QY 188 ccctgcgagcaacatctcatgcagaagatccaacgtgacgaggattcatatgaacgggac 247
|||||

Db 121 ccctgcgagcaacatctcatgcagaagatccaacgtgacgaggattcatatgaacgggac 180

QY 248 ccgtacagccctagtcaggatccgtacagccctagtcctcatatgatcggagagcgctgga 307
|||||

Db 181 ccgtacagccctagtcaggatccgtacagccctagtcctcatatgatcggagagcgctgga 240

QY 308 tcctctcagcaccacagagaggtgttgcaatgagctgaacgagtttgagaacaaccaaagg 367
|||||

Db 241 tcctctcagcaccacagagaggtgttgcaatgagctgaacgagtttgagaacaaccaaagg 300

QY 368 tgcattgctgcaggcattgcaacagatcatggagaaacagagcgataggttcaggaggagg 427
Db 301 tgcattgctgcaggcattgcaacagatcatggagaaacagagcgataggttcaggaggagg 360

QY 428 caacaggagcaacagattcaagaggagctcaggaactgcctcaacagtcggccttagg 487
Db 361 caacaggagcaacagattcaagaggagctcaggaactgcctcaacagtcggccttagg 420

QY 488 gcaccacagcgttgcgacttgacgtcgaaagtgcggcgag 528
Db 421 gcaccacagcgttgcgacttgacgtcgaaagtgcggcgag 461

RESULT 7
AAH45081/c
ID AAH45081 standard; DNA; 3150 BP.
XX AC AAH45081;
XX AC AAH45081;
DT 04-SEP-2001 (first entry)
DE lacZ gene.
XX lacZ; immunosuppressive; epigenetic regulation motif; immune response;
KW T-cell response; methylation activity;
KW methyl DNA binding protein identification; ds.
XX Escherichia coli.
OS
XX WO200140478-A2.
PN
XX 07-JUN-2001.
PD
XX
PF 06-DEC-2000; 2000WO-EP12793.
XX
PR 06-DEC-1999; 99CA-2291367.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CENT NAT RECH SCI.
XX
PI Nicolas J, Henry I, Choulika A;
XX WPI; 2001-367812/38.
DR
XX
PT New isolated polynucleotide having reduced or increased content of
PT epigenetic control motifs for studying, increasing and/or reducing gene
PT expression, and improving DNA vaccination methods -
XX Example 1; Fig 1; 75pp; English.
PS
XX The present invention relates to modified lacZ genes (see AAH45079 and
CC AAH45080), in which epigenetic regulation motifs have been mutated
CC compared to the wild-type sequence (the present sequence). The mutant
CC lacZ genes are useful for inducing in a second host, a protective immune
CC response, against a gene product of a first host. The mutant lac Z genes
CC are also useful for evaluating a promoter in biological systems, for
CC comparing methylation activity in biologically systems and/or for
CC identifying unknown methyl DNA binding proteins. The mutant lacZ genes
CC are also useful for compensating a genetic defect, and for therapeutic
CC applications. The mutant lacZ genes are also useful for minimising a
CC T-cell response against the T-cells or tissues treated with them.
XX
SQ Sequence 3150 BP; 706 A; 853 C; 904 G; 687 T; 0 other;

Query Match 21.4%; Score 146; DB 22; Length 3150;
Best Local Similarity 100.0%; Pred. No. 5.7e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggctccaggagtcgtcgccaccaatcccccatatggaacc 596
Db 3099 GAATTCCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCCACCAATCCCCCATATGGAACC 3040

QY 597 gtcgatattcagccatgtgcctttctccgcgtgcagcagatggcgatggctggtttccat 656
Db 3039 GTCGATATTACGCCATGTGCCCTTCTTCCGGTGCAGCAGATGGCGTGGTTTCCAT 2980

QY 657 cagttgctgtgactgtagcggctga 682
Db 2979 CAGTTGCTGTGACTGTAGCGGCTGA 2954

RESULT 8
AAQ41287/c
ID AAQ41287 standard; DNA; 3365 BP.
XX AC AAQ41287;
XX AC AAQ41287;
DT 26-AUG-1993 (first entry)
XX Ubiquitin fusion protein, Ub-mat-beta-Gal, coding sequence.
DE
XX Fusion protein; ubiquitin-methionine-beta-galactosidase; UBP2; UBP3;
KW expression vector; ubiquitin-specific; protease; UBP1; in vitro;
KW transformation; Ub-met-beta-Gal; in vivo; ss.
XX
OS Synthetic.
XX WO9309235-A.
PN
XX 13-MAY-1993.
PD
XX
PF 06-NOV-1992; 92WO-US09562.
XX
PR 08-NOV-1991; 91US-0789915.
XX
PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX Baker RT, Tobias JW, Varshavsky A;
PI
XX WPI; 1993-167700/20.
DR P-PSDB; AAR36728.
XX
PT Ubiquitin-specific protease(s) - specifically cleave fusion
PT proteins regardless of size, at C-terminus of ubiquitin portion
XX
PS Claim 2; Page 31-36; 78pp; English.
XX
CC This sequence encodes the 120 kD fusion protein ubiquitin-methionine-
CC beta-galactosidase (Ub-met-beta-Gal). This protein was used as a
CC substrate to demonstrate the specific cleavage of the ubiquitin-
CC specific proteases UBP1, UBP2, and UBP3 (see also AAR36729-31). The
CC proteases UBP1 and UBP2 demonstrate activity both in vivo and in
CC vitro, whereas UBP3 is only active in vivo. These proteases have
CC been shown to specifically cleave Ub-met-beta-Gal at the C-terminus
CC of the ubiquitin moiety. The proteases can be used to deubiquinate
CC fusion proteins in vivo, therefore prokaryotic cells having an
CC expression vector one of these proteases can be further transformed
CC with an expression vector encoding a ubiquitin fusion protein.
CC Such cells will then produce a deubiquinated product having a
CC predetermined N-terminal amino acid residue.
XX
SQ Sequence 3365 BP; 765 A; 917 C; 951 G; 732 T; 0 other;

Query Match 21.4%; Score 146; DB 14; Length 3365;
Best Local Similarity 100.0%; Pred. No. 5.9e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggctccaggagtcgtcgccaccaatcccccatatggaacc 596
Db 3315 GAATTCCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCCACCAATCCCCCATATGGAACC 3256

QY 597 gtcgatattcagccatgtgcctttctccgcgtgcagcagatggcgatggctggtttccat 656
Db 3255 GTCGATATTACGCCATGTGCCCTTCTTCCGGTGCAGCAGATGGCGATGGCTGGTTTCCAT 3196

QY 657 cagttgctgttgactgtagcggtga 682
|||||
Db 3195 CAGTTGCTGTTGACTGTAGCGGCTGA 3170

RESULT 9
AAD04924/c
ID AAD04924 standard; DNA; 3652 BP.
XX
AC AAD04924;
XX
DT 17-JUL-2001 (first entry)
XX
DE E. coli beta-galactosidase reporter gene of gene trapping construct.
XX
KW Gene trapping construct; conditional mutation; unidirectional inversion;
KW recombinase recognition sequence; RRS; disruption cassette;
KW selection cassette; transgenic organism; beta-galactosidase; ds.
XX
OS Escherichia coli.
XX
PN WO200129208-A1.
XX
PD 26-APR-2001.
XX
PF 16-OCT-2000; 2000WO-EP10162.
XX
PR 16-OCT-1999; 99EP-0120592.
PR 27-OCT-1999; 99US-0162016.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX
PI Kuehn R, Von Melchener H, Altschmied J;
XX
DR WPI; 2001-308486/32.
XX
PT New gene trapping construct capable of causing conditional mutations in
PT genes, comprises functional DNA segment inserted in sense or antisense
PT direction relative to gene to be trapped -
XX
PS Disclosure; Page 51-52; 78pp; English.
XX
CC The present invention relates to a conditional gene trapping construct
CC capable of causing conditional mutations in genes. The gene trapping
CC construct comprises two functional DNA segments, each being flanked by
CC two recombinase recognition sequences (RRSs) specific to site specific
CC recombinase which is capable of unidirectional inversion of double
CC standard DNA segment. One of the DNA segment (disruption cassette) is
CC inserted in antisense orientation relative to the transcriptional
CC orientation of the gene to be trapped. The other DNA segment (selection
CC cassette) is inserted in sense direction relative to the transcriptional
CC orientation of the gene to be trapped. The cell comprising the gene
CC trapping construct is useful for the identification and/or isolation of
CC genes. The transgenic organism comprising the gene trapping construct is
CC useful to study gene function at various developmental stages. The gene
CC trapping construct is useful for mutationally inactivating all cellular
CC genes. The present sequence is Escherichia coli beta-galactosidase
CC reporter gene of disruption cassette functional DNA segment of gene
CC trapping construct.
SQ Sequence 3652 BP; 830 A; 971 C; 1000 G; 851 T; 0 other;

Query Match 21.4%; Score 146; DB 22; Length 3652;
Best Local Similarity 100.0%; Pred. No. 6.1e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatccccataggaaacc 596
|||||
Db 3207 GAATTCCGCCGATACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCCATATGGAACC 3148

QY 597 gtcgatatccagcccatgtgcctttccgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 3147 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTCCAT 3088

QY 657 cagttgctgttgactgtagcggtga 682
|||||
Db 3087 CAGTTGCTGTTGACTGTAGCGGCTGA 3062

RESULT 10
AAF85070/c
ID AAF85070 standard; DNA; 3652 BP.
XX
AC AAF85070;
XX
DT 09-JUL-2001 (first entry)
XX
DE Nucleotide sequence of a DNA fragment from Escherichia coli.
XX
KW Gene trapping construct; conditional mutation; recombinase; ss.
XX
OS Escherichia coli.
XX
PN EP1092768-A1.
XX
PD 18-APR-2001.
XX
PF 16-OCT-1999; 99EP-0120592.
XX
PR 16-OCT-1999; 99EP-0120592.
XX
PA (ARTE-) ARTEMIS PHARM GMBH.
PA (FRAN-) FRANKGEN BIOTECHNOLOGIE AG.
XX
PI Kuehn R, Von Melchner H;
XX
DR WPI; 2001-275205/29.
XX
PT New gene trapping construct , useful for mutationally inactivating all
PT genes, comprises functional DNA segment inserted in antisense direction
PT relative to gene to be trapped -
XX
PS Disclosure; Page 11-12; 20pp; English.
XX
CC The specification describes a gene trapping construct which is
CC capable of causing conditional mutations in genes. The construct
CC comprises a functional DNA segment inserted in antisense direction
CC relative to the gene to be trapped, where the functional DNA segment
CC is flanked by two recombinase recognition sequences which are specific
CC to a site specific recombinase capable of inverting a double standard
CC DNA segment. The gene trapping construct is useful for identification
CC and isolation of genes. It may be used in methods to study gene function
CC at various developmental stages. The construct is also useful for
CC mutationally inactivating all cellular genes. The present sequence
CC represents a DNA fragment from Escherichia coli, which is used to
CC produce constructs of the invention.
XX
SQ Sequence 3652 BP; 830 A; 971 C; 1000 G; 851 T; 0 other;

Query Match 21.4%; Score 146; DB 22; Length 3652;
Best Local Similarity 100.0%; Pred. No. 6.1e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatccccataggaaacc 596
|||||
Db 3207 GAATTCCGCCGATACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCCATATGGAACC 3148

QY 597 gtcgatatccagcccatgtgcctttccgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 3147 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTCCAT 3088

QY 657 cagttgctgttgactgtagcggtga 682

```
Db 3087 CAGTGTCTGTTGACTGTAGCGGCTGA 3062
|||||
RESULT 11
AAX87978/c
ID AAX87978 standard; DNA; 3756 BP.
XX
AC AAX87978;
XX
DT 06-DEC-1999 (first entry)
XX
DE HCMV UL45 promoter:lac2 gene fusion.
XX
KW HCMV; UL45 gene; promoter; lac2 gene; infection; diagnosis;
KW herpesvirus; ss.
XX
OS Chimeric - Human cytomegalovirus.
OS Chimeric - Escherichia coli.
XX
FH Key Location/Qualifiers
FT promoter 1..483
FT /*tag= a
FT /note= "UL45 gene promoter"
FT CDS 484..3714
FT /*tag= b
XX
PN US9598676-A.
XX
PD 28-SEP-1999.
XX
PF 25-APR-1997; 97US-0846026.
XX
PR 18-JUN-1992; 92US-0900279.
PR 28-FEB-1995; 95US-0395673.
XX
PA (UNIW ) UNIV WASHINGTON.
XX
PI Olivo PD;
XX
DR WPI; 1999-561024/47.
DR P-PSDB; AAY31768.
XX
PT Genetically engineered cell lines for detecting infectious Herpesvirus
PT -
XX
PS Example 17; Fig 13; 44pp; English.
XX
CC This is the nucleotide sequence of a UL45 promoter/reporter gene
CC construct composed of the human cytomegalovirus (HCMV) UL45 gene
CC promoter and 177 bp of the UL45 open reading frame fused to the
CC Escherichia coli LacZ gene in pCMVUL45LacZ. The predicted product
CC (see AAY31768) of the chimeric gene is a fusion protein containing 59
CC amino acids of the UL45 protein fused to the N-terminus of
CC beta-galactosidase. A claimed in vitro diagnostic method for
CC detecting infectious HCMV in a specimen comprises: (a) providing a
CC genetically engineered cell line (especially a mink lung cell line)
CC stably transformed with DNA containing a beta-gene promoter
CC sequence of HCMV (especially the UL45 promoter) linked to a
CC reporter gene (especially a beta-galactosidase gene), whose
CC expression is dependent upon the presence of HCMV; (b) inoculating
CC the cell line with a specimen suspected of containing HCMV; (c)
CC allowing the infectious cycle of the virus to proceed; and (d)
CC detecting the infectious virus, especially by detecting
CC beta-galactosidase activity on a fluorogenic substrate or by
CC histochemical/light microscopy.
XX
SQ Sequence 3756 BP; 836 A; 1033 C; 1079 G; 808 T; 0 other;

Query Match 21.4%; Score 146; DB 20; Length 3756;
Best Local Similarity 100.0%; Pred. No. 6.2e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 537 gaattccgcgatactgacgggtccaggagtcgtagcagcaccatccccataggaaacc 596
|||||
Db 3663 GAATTCGCGGATACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCATATGGAAC 3604
|||||
QY 597 gtcgatatccagccatgtgccttcttcctccgctgcagcagatggcgatggctgttccat 656
|||||
Db 3603 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTCCAT 3544
|||||
QY 657 cagttgctgttgactgtagcggtga 682
|||||
Db 3543 CAGTTGCTGTTGACTGTAGCGGCTGA 3518
|||||

RESULT 12
AAX06800/c
ID AAX06800 standard; DNA; 4279 BP.
XX
AC AAX06800;
XX
DT 26-APR-1999 (first entry)
XX
DE Multiple cloning site/Kozak sequence/Lac2 gene construct.
XX
KW Gene activation; gene expression; cancer; antisense; therapy;
KW beta-galactosidase; luciferase; ss.
XX
OS Chimeric - Escherichia coli.
OS Chimeric - synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1..64
FT /*tag= a
FT /function= multiple cloning site
FT misc_signal 65..79
FT /*tag= b
FT /function= Kozak sequence consensus
FT prim_transcript 80..4279
FT /*tag= c
FT /standard_name= beta-galactosidase
XX
PN WO9858944-A1.
XX
PD 30-DEC-1998.
XX
PF 24-JUN-1998; 98WO-US13093.
XX
PR 25-JUN-1997; 97US-0050772.
XX
PA (BLAC/) BLACK C A.
XX
PI Black CA;
XX
DR WPI; 1999-095360/08.
XX
PT New masked targetted expression cassette - useful for regulating
PT gene expression, and preventing neoplastic cell growth
XX
PS Disclosure; Page 22-24; 43pp; English.
XX
CC This is the multiple cloning site-Kozak-lac2 gene sequence of
CC armed sense strand plasmid pCI-Neo, which was designed for the
CC production of the sense RNA strand of a targetted expression
CC cassette. Alternative flanking sequences (see AAX06801-07)
CC corresponding to portions of the firefly luciferase mRNA are
CC inserted into the multiple cloning sites such that transcription
CC from the T7 promoter yields RNA comprising (from the 5' end):
CC luciferase segment-Kozak-beta-galactosidase. Sense strand RNA of
CC a masked targetted cassette is produced by in vitro transcription
CC of the construct. Antisense sequences (see AAX06808-14)
CC corresponding to portions of the target molecule are hybridised to
CC complementary flanking sequences of the sense strand of the
CC targetted cassette. The hybridised mixture is introduced to an in
```


PT into cancer killing drug
XX
PS Claim 4; Fig 19B; 178pp; English.
XX
CC This sequence comprises Region A of a replication-deficient adenovirus
CC type 5 vector containing a lacZ gene (encoding beta-galactosidase
CC (beta-gal)) under the control of the rat probasin promoter.
CC Inducing cellular cytotoxicity of a tumor cell comprises introducing a
CC replication-deficient adenovirus type 5 expression vector comprising a
CC gene that encodes for an enzyme that has the ability to convert a
CC non-toxic pro-drug into a cancer killing drug which then destroys cancer
CC cells. The adenovirus genome preferably has a deletion in an E1 and E3
CC region and an insertion within the region of a nucleic acid encoding
CC Escherichia coli beta-gal under the control of a
CC promoter. The pro-drug active site is masked by beta-gal. Functional
CC beta-gal is expressed from the vector so as to activate the pro-drug into
CC an agent toxic to the cells. Beta-gal can be under the control of a
CC Rous Sarcoma Virus (RSV), probasin (PB), Prostate Specific Antigen (PSA)
CC or Mouse Mammary Tumor Virus (MMTV) promoter. The vectors provide a
CC novel way to treat prostate cancer by gene therapy.
XX
SQ Sequence 5201 BP; 1318 A; 1248 C; 1345 G; 1290 T; 0 other;

Query Match 21.4%; Score 146; DB 21; Length 5201;
Best Local Similarity 100.0%; Pred. No. 7e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgcgatactgacgggctccaggagtcgtcgccaccaatcccccatatggaaacc 596
|||||
Db 4057 GAATTCCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCCACCAATCCCCATATGGAAACC 3998

QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 3997 GTCGATATTACGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTTCCAT 3938

QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 3937 CAGTTGCTGTTGACTGTAGCGGCTGA 3912

RESULT 15
AAA09089/c
ID AAA09089 standard; DNA; 5346 BP.
XX
AC AAA09089;
XX
DT 10-AUG-2000 (first entry)
XX
DE AdPSA-beta-galactosidase vector region A.
XX
KW Region A; replication-deficient; vector; lacZ; beta-galactosidase; PSA;
KW prostate specific antigen; promoter; cytotoxicity; cytostatic; pro-drug;
KW prostate cancer; gene therapy; ss.
XX
OS Chimeric - Adenovirus type 5.
OS Chimeric - Homo sapiens.
OS Chimeric - Escherichia coli.
XX
PN WO200020038-A1.
XX
PD 13-APR-2000.
XX
PF 01-OCT-1999; 99WO-US20907.
XX
PR 02-OCT-1998; 98US-0165730.
XX
PA (GENO-) GENOTHERAPEUTICS INC.
XX
PI Steiner MS;
XX
DR WPI; 2000-303646/26.
XX

PT Inducing cellular cytotoxicity of tumor cell comprises introducing
PT replication-deficient adenovirus type 5 expression vector containing
PT gene encoding for enzyme having ability to convert nontoxic prodrug
PT into cancer killing drug
XX
PS Claim 4; Fig 20B; 178pp; English.
XX
CC This sequence comprises Region A of a replication-deficient adenovirus
CC type 5 vector containing a lacZ gene (encoding beta-galactosidase
CC (beta-gal)) under the control of the prostate specific antigen (PSA)
CC promoter.
CC Inducing cellular cytotoxicity of a tumor cell comprises introducing a
CC replication-deficient adenovirus type 5 expression vector comprising a
CC gene that encodes for an enzyme that has the ability to convert a
CC non-toxic pro-drug into a cancer killing drug which then destroys cancer
CC cells. The adenovirus genome preferably has a deletion in an E1 and E3
CC region and an insertion within the region of a nucleic acid encoding
CC Escherichia coli beta-gal under the control of a
CC promoter. The pro-drug active site is masked by beta-gal. Functional
CC beta-gal is expressed from the vector so as to activate the pro-drug into
CC an agent toxic to the cells. Beta-gal can be under the control of a
CC Rous Sarcoma Virus (RSV), probasin (PB), PSA or Mouse Mammary Tumor Virus
CC (MMTV) promoter. The vectors provide a novel way to treat prostate
CC cancer by gene therapy.
XX
SQ Sequence 5346 BP; 1295 A; 1283 C; 1447 G; 1321 T; 0 other;

Query Match 21.4%; Score 146; DB 21; Length 5346;
Best Local Similarity 100.0%; Pred. No. 7.1e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgcgatactgacgggctccaggagtcgtcgccaccaatcccccatatggaaacc 596
|||||
Db 4204 GAATTCCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCCACCAATCCCCATATGGAAACC 4145

QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 4144 GTCGATATTACGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTTCCAT 4085

QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 4084 CAGTTGCTGTTGACTGTAGCGGCTGA 4059

Search completed: December 27, 2001, 13:14:57
Job time: 14473 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 13:06:53 ; Search time 178.26 Seconds
(without alignments)
866.476 Million cell updates/sec

Title: US-09-715-036-3
Perfect score: 682
Sequence: 1 gacacagaccaactggtaat.....ctgttgacttagcggtga 682

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| c 1 | 146 | 21.4 | 3152 | 1 US-07-924-028A-3 | Sequence 3, Appli |
| c 2 | 146 | 21.4 | 3365 | 1 US-07-789-915A-1 | Sequence 1, Appli |
| c 3 | 146 | 21.4 | 3365 | 1 US-08-005-002C-1 | Sequence 1, Appli |
| c 4 | 146 | 21.4 | 3365 | 1 US-08-487-203A-1 | Sequence 1, Appli |
| 5 | 146 | 21.4 | 8710 | 1 US-08-480-882B-3 | Sequence 3, Appli |
| 6 | 146 | 21.4 | 8710 | 1 US-08-480-210-3 | Sequence 3, Appli |
| 7 | 146 | 21.4 | 9019 | 1 US-08-480-882B-4 | Sequence 4, Appli |
| 8 | 146 | 21.4 | 9019 | 1 US-08-480-210-4 | Sequence 4, Appli |
| 9 | 146 | 21.4 | 9641 | 2 US-08-374-483-3 | Sequence 4, Appli |
| 10 | 146 | 21.4 | 10241 | 1 US-08-508-448C-16 | Sequence 16, Appl |
| 11 | 146 | 21.4 | 10281 | 1 US-08-816-155B-1 | Sequence 1, Appli |
| 12 | 146 | 21.4 | 10281 | 3 US-09-079-587-1 | Sequence 1, Appli |
| 13 | 146 | 21.4 | 34382 | 2 US-08-374-483-6 | Sequence 6, Appli |
| c 14 | 145 | 21.3 | 7852 | 3 US-08-836-022A-2 | Sequence 2, Appli |
| c 15 | 145 | 21.3 | 7852 | 4 US-09-427-048A-2 | Sequence 2, Appli |
| c 16 | 145 | 21.3 | 8313 | 1 US-08-232-463-2 | Sequence 2, Appli |
| c 17 | 145 | 21.3 | 8509 | 1 US-08-462-014-1 | Sequence 1, Appli |
| c 18 | 145 | 21.3 | 8509 | 4 US-08-973-334-4 | Sequence 4, Appli |
| c 19 | 145 | 21.3 | 8509 | 4 US-09-563-869A-4 | Sequence 4, Appli |
| 20 | 145 | 21.3 | 8775 | 1 US-08-232-463-5 | Sequence 5, Appli |
| c 21 | 145 | 21.3 | 9454 | 1 US-08-232-463-3 | Sequence 3, Appli |
| c 22 | 145 | 21.3 | 9454 | 1 US-08-232-463-4 | Sequence 4, Appli |
| c 23 | 145 | 21.3 | 9661 | 3 US-08-716-351A-3 | Sequence 3, Appli |
| 24 | 145 | 21.3 | 9890 | 1 US-08-232-463-18 | Sequence 18, Appl |
| 25 | 145 | 21.3 | 9916 | 1 US-08-232-463-17 | Sequence 17, Appl |
| 26 | 145 | 21.3 | 9917 | 1 US-08-232-463-16 | Sequence 16, Appl |
| c 27 | 145 | 21.3 | 10306 | 3 US-08-716-351A-4 | Sequence 4, Appli |

| | | | | | |
|------|-------|------|-------|--------------------|-------------------|
| c 28 | 145 | 21.3 | 10398 | 2 US-08-331-384-1 | Sequence 1, Appli |
| c 29 | 145 | 21.3 | 10398 | 2 US-08-708-188-1 | Sequence 1, Appli |
| c 30 | 145 | 21.3 | 10398 | 2 US-08-836-087-1 | Sequence 1, Appli |
| c 31 | 145 | 21.3 | 10398 | 4 US-09-246-320-1 | Sequence 1, Appli |
| c 32 | 145 | 21.3 | 10398 | 4 US-09-242-743-1 | Sequence 1, Appli |
| 33 | 145 | 21.3 | 10408 | 1 US-08-232-463-6 | Sequence 6, Appli |
| c 34 | 145 | 21.3 | 10408 | 1 US-08-232-463-7 | Sequence 7, Appli |
| c 35 | 145 | 21.3 | 10970 | 3 US-08-716-351A-5 | Sequence 5, Appli |
| c 36 | 145 | 21.3 | 35408 | 4 US-08-973-334-3 | Sequence 3, Appli |
| c 37 | 145 | 21.3 | 35408 | 4 US-09-563-869A-3 | Sequence 3, Appli |
| c 38 | 145 | 21.3 | 35408 | 4 US-08-549-489-3 | Sequence 3, Appli |
| 39 | 143.4 | 21.0 | 7897 | 3 US-08-836-022A-1 | Sequence 1, Appli |
| 40 | 143.4 | 21.0 | 7897 | 4 US-09-427-048A-1 | Sequence 1, Appli |
| c 41 | 142.8 | 20.9 | 4810 | 3 US-08-852-629-11 | Sequence 11, Appl |
| c 42 | 142.8 | 20.9 | 4838 | 3 US-08-852-629-15 | Sequence 15, Appl |
| c 43 | 142.8 | 20.9 | 19182 | 2 US-08-850-880-11 | Sequence 11, Appl |
| c 44 | 142.8 | 20.9 | 19182 | 2 US-08-944-916-11 | Sequence 11, Appl |
| 45 | 140.2 | 20.6 | 619 | 5 PCT-US95-13658-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-07-924-028A-3/c
; Sequence 3, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3152 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-924-028A-3

Query Match 21.4%; Score 146; DB 1; Length 3152;
Best Local Similarity 100.0%; Pred. No. 1.6e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatcccccatatggaaacc 596
|||||
Db 3036 GAATTCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCACCAATCCCCATATGGAACC 2977

Qy 597 gtcgatattcagccatgtgccttcttcgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 2976 GTCGATATTCAGCCATGTGCCTTCTTCCGGTGCAGCAGATGGCGATGGCTGGTTCCAT 2917

Qy 657 cagttgctgttgactgtagcggctga 682
|||||
Db 2916 CAGTTGCTGTGACTGTAGCGGCTGA 2891

RESULT 2
US-07-789-915A-1/c
; Sequence 1, Application US/07789915A
; Patent No. 5212058
; GENERAL INFORMATION:
; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.
; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,915A
; FILING DATE: 19911108
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-5091AA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3365 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3363
US-07-789-915A-1

Query Match 21.4%; Score 146; DB 1; Length 3365;
Best Local Similarity 100.0%; Pred. No. 1.6e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatcccccatatggaaacc 596
|||||
Db 3315 GAATTCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCACCAATCCCCATATGGAACC 3256

Qy 597 gtcgatattcagccatgtgccttcttcgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 3255 GTCGATATTCAGCCATGTGCCTTCTTCCGGTGCAGCAGATGGCGATGGCTGGTTCCAT 3196

Qy 657 cagttgctgttgactgtagcggctga 682
|||||

Db 3195 CAGTTGCTGTGACTGTAGCGGCTGA 3170

RESULT 3
US-08-005-002C-1/c
; Sequence 1, Application US/08005002C
; Patent No. 5494818
; GENERAL INFORMATION:
; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.
; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: Maine
; COUNTRY: U.S.A.
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/005,002C
; FILING DATE: 15-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,915
; FILING DATE: 08-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: MIT-5091AAZ
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 207-363-0558
; TELEFAX: 207-363-0528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3363
US-08-005-002C-1

Query Match 21.4%; Score 146; DB 1; Length 3365;
Best Local Similarity 100.0%; Pred. No. 1.6e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatcccccatatggaaacc 596
|||||
Db 3315 GAATTCGCCGATACTGACGGGCTCCAGGAGTCGTGCGCACCAATCCCCATATGGAACC 3256

Qy 597 gtcgatattcagccatgtgccttcttcgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 3255 GTCGATATTCAGCCATGTGCCTTCTTCCGGTGCAGCAGATGGCGATGGCTGGTTCCAT 3196

Qy 657 cagttgctgttgactgtagcggctga 682
|||||
Db 3195 CAGTTGCTGTGACTGTAGCGGCTGA 3170

RESULT 4
US-08-487-203A-1/c
; Sequence 1, Application US/08487203A
; Patent No. 5683904
; GENERAL INFORMATION:

; APPLICANT: Baker, Rohan T.
; APPLICANT: Tobias, John W.
; APPLICANT: Varshavsky, Alexander
; TITLE OF INVENTION: Ubiquitin-Specific Proteases
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: Maine
; COUNTRY: U.S.A.
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,203A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,002
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: MIT-5091A3Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 207-363-0558
; TELEFAX: 207-363-0528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3363
US-08-487-203A-1

Query Match 21.4%; Score 146; DB 1; Length 3365;
Best Local Similarity 100.0%; Pred. No. 1.6e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatccccatatggaacc 596
|||||
Db 3315 GAATTCGCCGACTGACGGGCTCCAGAGTCGTGCCACCAATCCCATATGGAAC 3256
|||||

QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 3255 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTTCCAT 3196
|||||

QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 3195 CAGTTGCTGTTGACTGTAGCGGCTGA 3170
|||||

RESULT 5
US-08-480-882B-3
; Sequence 3, Application US/08480882B
; Patent No. 5656275
; GENERAL INFORMATION:
; APPLICANT: WASMOEN, TERRI
; APPLICANT: CHAVEZ, LLOYD
; APPLICANT: CHU, HSIEH-JUE
; TITLE OF INVENTION: RECOMBINANT RACCOON POX VIRUSES
; TITLE OF INVENTION: AND THEIR USE AS AN EFFECTIVE VACCINE AGAINST FELINE
; TITLE OF INVENTION: INFECTIOUS PERITONITIS VIRUS DISEASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby, P.C.
; STREET: 805 Third Avenue 27th Floor
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,882B
; FILING DATE: 07-JUN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,516
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 0632/08669
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2125277700
; TELEFAX: 2127536237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Feline infectious peritonitis virus
; IMMEDIATE SOURCE:
; CLONE: psc11f1
US-08-480-882B-3

Query Match 21.4%; Score 146; DB 1; Length 8710;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgccgatactgacgggtccaggagtcgtcgccaccaatccccatatggaacc 596
|||||
Db 3102 GAATTCGCCGACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCATATGGAAC 3161
|||||

QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 3162 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGTTTCCAT 3221
|||||

QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 3222 CAGTTGCTGTTGACTGTAGCGGCTGA 3247
|||||

RESULT 6
US-08-480-210-3
; Sequence 3, Application US/08480210
; Patent No. 5770211
; GENERAL INFORMATION:
; APPLICANT: WASMOEN, TERRI
; APPLICANT: CHAVEZ, LLOYD
; APPLICANT: CHU, HSIEH-JUE
; TITLE OF INVENTION: RECOMBINANT RACCOON POX VIRUSES
; TITLE OF INVENTION: AND THEIR USE AS AN EFFECTIVE VACCINE AGAINST FELINE
; TITLE OF INVENTION: INFECTIOUS PERITONITIS VIRUS DISEASE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby, P.C.
; STREET: 805 Third Avenue 27th Floor
; CITY: New York

STATE: NY
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,210
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/125,516
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 0632/18669-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2125277700
TELEFAX: 2127536237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 8710 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Feline infectious peritonitis virus
IMMEDIATE SOURCE:
CLONE: psc11f1
US-08-480-210-3

Query Match 21.4%; Score 146; DB 1; Length 8710;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 537 gaattccgcgatactgacgggctccaggagtcgtcgccaccaatcccccatatggaaaacc 596
|||||
Db 3102 GAATTCGGCCGATACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCCATATGGAAACC 3161
QY 597 gtcgatattcagccatgtgccttcttcgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 3162 GTCGATATTCAGCCATGTGCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTTCCAT 3221
QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 3222 CAGTTGCTGTGACTGTAGCGGCTGA 3247

RESULT 7
US-08-480-882B-4
Sequence 4, Application US/08480882B
Patent No. 5656275
GENERAL INFORMATION:
APPLICANT: WASMOEN, TERRI
APPLICANT: CHAVEZ, LLOYD
APPLICANT: CHU, HSIEN-JUE
TITLE OF INVENTION: RECOMBINANT RACCOON POX VIRUSES
TITLE OF INVENTION: AND THEIR USE AS AN EFFECTIVE VACCINE AGAINST FELINE
TITLE OF INVENTION: INFECTIOUS PERITONITIS VIRUS DISEASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue 27th Floor
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,882B
FILING DATE: 07-JUN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/125,516
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 0632/08669
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2125277700
TELEFAX: 2127536237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 9019 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Feline immunodeficiency virus
IMMEDIATE SOURCE:
CLONE: psc11e1
US-08-480-882B-4

Query Match 21.4%; Score 146; DB 1; Length 9019;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 537 gaattccgcgatactgacgggctccaggagtcgtcgccaccaatcccccatatggaaaacc 596
|||||
Db 3101 GAATTCGGCCGATACTGACGGGCTCCAGGAGTCGTGCCACCAATCCCCATATGGAAACC 3160
QY 597 gtcgatattcagccatgtgccttcttcgcgctgcagcagatggcgatggctggtttccat 656
|||||
Db 3161 GTCGATATTCAGCCATGTGCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTTCCAT 3220
QY 657 cagttgctgttgactgtagcggctga 682
|||||
Db 3221 CAGTTGCTGTGACTGTAGCGGCTGA 3246

RESULT 8
US-08-480-210-4
Sequence 4, Application US/08480210
Patent No. 5770211
GENERAL INFORMATION:
APPLICANT: WASMOEN, TERRI
APPLICANT: CHAVEZ, LLOYD
APPLICANT: CHU, HSIEN-JUE
TITLE OF INVENTION: RECOMBINANT RACCOON POX VIRUSES
TITLE OF INVENTION: AND THEIR USE AS AN EFFECTIVE VACCINE AGAINST FELINE
TITLE OF INVENTION: INFECTIOUS PERITONITIS VIRUS DISEASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby, P.C.
STREET: 805 Third Avenue 27th Floor
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,210
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/125,516
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 0632/18669-US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 2125277700
TELEFAX: 2127536237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 9019 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Feline immunodeficiency virus
IMMEDIATE SOURCE:
CLONE: psc11e1
US-08-480-210-4

Query Match 21.4%; Score 146; DB 1; Length 9019;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgcgatactgacgggctccaggagtcgcccacccaatcccccatatggaaaacc 596
|||||
Db 3101 GAATTCGCCGATACTGACGGGCTCCAGGAGTCGTCGCCACCAATCCCCATATGGAAACC 3160
QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 3161 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTTCCAT 3220
QY 657 cagttgctgttgactgtagcggtga 682
|||||
Db 3221 CAGTTGCTGTTGACTGTAGCGGCTGA 3246

RESULT 9
US-08-374-483-3
Sequence 3, Application US/08374483
Patent No. 5880102
GENERAL INFORMATION:
APPLICANT: GEORGE, SAMUEL E.
APPLICANT: BLAZING, MICHAEL A.
TITLE OF INVENTION: ADENOVIRAL VECTOR SYSTEM
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,483
FILING DATE: 17-JAN-1995
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-83
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9641 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-374-483-3

Query Match 21.4%; Score 146; DB 2; Length 9641;
Best Local Similarity 100.0%; Pred. No. 2.5e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 537 gaattccgcgatactgacgggctccaggagtcgcccacccaatcccccatatggaaaacc 596
|||||
Db 464 GAATTCGCCGATACTGACGGGCTCCAGGAGTCGTCGCCACCAATCCCCATATGGAAACC 523
QY 597 gtcgatattcagccatgtgccttcttcgcggtgcagcagatggcgatggctggtttccat 656
|||||
Db 524 GTCGATATTCAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTTCCAT 583
QY 657 cagttgctgttgactgtagcggtga 682
|||||
Db 584 CAGTTGCTGTTGACTGTAGCGGCTGA 609

RESULT 10
US-08-508-448C-16
Sequence 16, Application US/08508448C
Patent No. 5804410
GENERAL INFORMATION:
APPLICANT: Kazuyoshi YAMAOKA et al.
TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING
TITLE OF INVENTION: TRYPSIN-LIKE ENZYME AND PROCESS FOR PRODUCING THE ENZYME
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/508,448C
FILING DATE: July 28, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 10241 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: cyclic
; MOLECULE TYPE: other nucleic acid (vector DNA)
US-08-508-448C-16

```

```
Query Match      21.4%; Score 146; DB 1; Length 10241;
Best Local Similarity 100.0%; Pred. NO. 2.5e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

| | | | |
|-----------|------|--|------|
| Qy | 537 | gaattccgccgatactacaggggctccaggagtcgtcgccaccaatccccatatggaacc | 596 |
| | | | |
| Db | 2426 | GAATTCCGCCGATACTGACCGGGCTCCAGGAGTCGTGCGCACCAATCCCATATGGAAACC | 2485 |

| | | | |
|-----------|------|---|------|
| Oy | 597 | gtcgatattcagcccatgtgccttcttcgcgtagcagcagatggcgatggctggtttccat | 656 |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| Dd | 2486 | GTCGATATTcAGCCCATGTCCTTCTTCGGCGTCACGCAGATGGCGATGGCTGTGTTTCCAT | 2545 |

Qy 657 cagttgctgttgactgtagcggctga 682
|||||
Db 2546 CAGTTGCTGTTGACTGTAGCGGCTGA 2571

```

RESULT      11
US-08-816-155B-1
; Sequence 1, Application US/08816155B
; Patent No. 5990091
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF

```

Query Match 21.4%; Score 146; DB 2; Length 10281;

```
Best Local Similarity 100.0%; Pred. No. 2.5e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 537 gaattccgcgatactgacggggtccaggagtcgtcgccaccaatccccatatggaacc 596
|||||
Db 7659 GAATTCGCGCGATACTGACGGGCTCCAGGAGTCGTGCGCCACCAATCCCCATATGGAACC 7718

Qy 597 gtcgatatccagccatgtgcctctctccgcgtgcagcagatggcgatggctggtttccat 656
|||
Db 7719 GTCGATATTCAGCCATGTGCCTTCTTCCGGTGCAGCAGATGGCGATGGCTGGTTTCCAT 7778

Qy 657 cagttgctgttgactgtagcgctga 682
|||||
Db 7779 CAGTTGCTGTTGACTGTAGCGGCTGA 7804

```

RESULT      12
US-09-079-587-1
; Sequence 1, Application US/09079587
; Patent No. 6130066
; GENERAL INFORMATION:
; APPLICANT: TARTAGLIA, JAMES
; APPLICANT: COX, WILLIAM I.
; APPLICANT: GETTIG, RUSSELL R.
; APPLICANT: MARTINEZ, HECTOR
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: PINCUS, STEVEN E.
; TITLE OF INVENTION: VECTORS HAVING ENHANCED EXPRESSION, AND
; TITLE OF INVENTION: METHODS OF MAKING AND USES THEREOF
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:

```

```
Query Match      21.4%; Score 146; DB 3; Length 10281;
Best Local Similarity 100.0%; Pred. No. 2.5e-33;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 537 gaattccgccgatactgacggggtccaggagtcgtcgccaccaatcccccatatggaaacc 596
|||||
Db 7659 GAATTCGCCGCGATACTACGGGGTCCAGGAGTCGTGCGCCACCAATCCCCATATGGAAC 7718


```

; Methods of Use Thereof
;
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/427,048A
; FILING DATE: 21-Oct-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,022
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GNVPN.008PCT
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
;
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-427-048A-2
```

```

Query Match      21.3%; Score 145; DB 4; Length 7852;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 538 aattccgccgatactgacggggtccaggagtcgcccaccaatccccatatggaaaccg 597
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4907 AATTCCGCCGATACTGACGGGCTCCAGGAGTCGTCGCCACCAATCCCCATATGGAAACCG 4848

Qy 598 tcgatattcagccatgtgccttcttcgcgtgcagcagatggcgatggctggttccatc 657
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4847 TCGATATTGAGCCATGTGCCTTCTTCCGCGTGCAGCAGATGGCGATGGCTGGTTCCATC 4788

Qy 658 agttgctgttgactgtagcggctga 682
      |||||||||||||||||||||||||||
Db 4787 AGTTGCTGTTGACTGTAGCGGCTGA 4763
```

Search completed: December 27, 2001, 13:07:22
Job time: 14058 sec

| | |
|----|---|
| XX | |
| PT | Producing transgenic peanut plants that produce allergen-free seeds, |
| PT | useful in non-allergenic foods, by antisense or sense co-suppression of |
| PT | allergen-encoding genes - |

PS Claim 20; Fig 4; 72pp; English.

CC The present sequence is that of peanut allergen Ara h3 cDNA. A
CC portion of the Ara h2 gene coding region is homologous to the
CC corresponding region of the peanut allergen Ara h4 gene. This
CC region has been PCR amplified, cloned in transformation vectors
CC (pUC18 and pBI4434) in sense and antisense orientations and used
CC to down-regulate Ara h3 and Ara h4 allergens in peanut. This is
CC an example of the method of the invention, which relates to the
CC production of a peanut plant having reduced, or undetectable,
CC allergenic protein (AP) content in its seed. A peanut plant cell
CC is transformed with a DNA construct containing an antisense AP
CC gene and/or sense AP gene, or their fragments, regenerated to
CC plants, and fertile transgenic plants that produce seeds with
CC reduced AP content are identified. The seeds are useful for
CC preparation of allergen-free foods.

Sequence 1855 BP; 564 A; 430 C; 465 G; 396 T; 0 other;

| | | | | |
|----------------------------|---------|--------------|--------|--------------|
| Query Match | 100.0%; | Score 1853; | DB 22; | Length 1855; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 1853; Conservative | 0; | Mismatches | 0; | Indels 0; |
| | | Gaps | 0; | |

| | |
|----|---|
| Qy | 1 atgsgtaagcttccttgagccttccctttgtccttgcctctacagtctctggagctaagcagc 60 |
| Db | 1 atgsgtaagcttccttgagccttccctttgtccttgcctctacagtctctggagctaagcagc 60 |

| | | |
|-----------|---|-----|
| QY | 61 atctccttcagcgacgaccggaaaggagaatgcgttgccagtccagcgctcaatgcgcag | 120 |
| | | |
| Dd | 61 atctccttcagcgacgaccggaaaggagaatgcgttgccagtccagcgctcaatgcgcag | 120 |

Qy 121 agacttgacaaaccgcattgaaatcgagggcggttacattgagacttgaaccccaacaac 180
|||||
Db 121 agacctgacaaccgcattgaaatcgagggcggttacattgagacttgaaccccaacaac 180

QY 181 caggagttcgaatgcgcgcgcgcgtcgcacctctctgcgttagtcctcgcgcgaacgcctt 240
|||||
Db 181 caggagttcgaatgcgcgcgcgcgtcgcacctctctgcgttagtcctcgcgcgaacgcctt 240

241 cgtagagcctttactccaatgctccccagagatcttcattccagcaagaaggatatac 300
 |||||
 241 cgtagagcctttactccaatgctccccagagatcttcattccagcaagaaggatatac 300
 |||||

301 ttgggttgatattccctgtgttcctagcacatatgaagagcctgcacacaagagcgc 360
301 ttgggttgatattccctgtgttcctagcacatatgaagagcctgcacacaagagcgc 360

361 cgataatcagtcaccaagaccaccaagagcgtttgccaagaagaagaccacaacgccaacagcaa 420
|||||
361 cgataatcagtcaccaagaccaccaagagcgtttgccaagaagaagaccacaacgccaacagcaa 420
|||||

421 caagatagtcaccaggaaggtgcacccgtttcaatgaggtgatctcattgcattccacc 480

481 ggtgttgcttctgtgctgacacgacacactgatgtgttgcgttctcttact 540

541 gacaccacaacaacagacaccagcttgatcagttccccaagagattcaatttggctgg 600

601 aaccagagcaagagttcttaaggtaccagcaacaagcagacgaagaagc 660
|||||
601 aaccacgagcaagagttcttaaggtaccagcaacaagcagacgaagaagc 660

| | | | |
|----|-----|--|-----|
| QY | 661 | ttaaccatatagcccatacagcccgcatagtcggcgctagaacgaaagagcgltgaattcgc | 720 |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| Db | 661 | ttaccatatagcccatacagcccgcatagtcggcgctagaacgaaagagcgltgaattcgc | 720 |

Dy 721 cctcgaagacagcacagccgcagaagaaacgagcaagaagaagggttga 780
|||||
bB 721 cctcgaggaacagcacagccgcagaagaaacgagcaagaagaagaagtcgtga 780

QY /81 aaacatctcagcgcttcacgcccggagttcctctgaaacaagccttcaggttcgacgacaga 840
 |||||
 Db 781 aaacatctcagcgcttcacgcccggagttcctctgaaacaagccttcaggttcgacgacaga 840

Oy 841 cagatctgcaaatctgtggygcgagacgagtgtaagaagaggagccattgtacg 900
 |||||
 Db 841 cagattgtgcaaaatctgtggygcgagacgagtgtaagaagaggagccattgtacg 900

[illegible]

961 gaatacgtagaagatcaatatgaataccacgaacagatgtgaagcgctggcaggggaagc 1020
 db 961 |||||
 961 gaatacgtagaagatcaatatgaataccacgaacagatgtgaagcgctggcaggggaagc 1020
 961 |||||

27 1021 agagcgcggggaatggtattgaagagacgactgcaccgcgatgtgttaaaagaacatt 1080
 |||
 |||
 Db 1021 agagcgcggggaatggtattgaagagacgactgcaccgcgatgtgttaaaagaacatt 1080

1001 ggcggaacacgacatccccacacatctacgacatcagcgctggttcactcaaaactgcccac 1140
 |||||
 1081 ggtggaacagatccccctacatctacgacatccctcagcgctggttcactcaaaactgcccac 1140

1141 gatctcaaccttaatccctagtgctgacttagtgctgataatggaatctctac 1200

1201 aggaatgcatgtttgtccctcactacacaccacaacgcacacgcatcatatatgtcatctg 1260

Db 1261 aggggacgycctcacgtgcaagtgtgagacagcaacgcaacagatgtatcagcagagag 1320

Db 1321 ctcaagaggtcacgttctgtgtccacagaactcgccgtgtggaagttccag 1380

Db 1381 agcgagaacttcgatactgtgcatcagaacagatltcaagccagcatagccaacttt 1440

Db 1441 gccggtgtaaacctcctcatagataaacctgccgaggggtgtgtgcaaatlcatatg c 1500

Db 1501 ctcccaaggagcagcgagcttaagaacaacaacccctcaagtctctgttcca 1560

Db 1561 cctttcagcagtcctccgagggctgtgcttaaaaaacgaccagtatcttttgcagcgctg 1620

Db 1621 ttatccactaacaataactttttgccacaatatgaataataataataagaagaataatgt 1680

QY 1681 agtttaatttttagtagaataagaatacacaagggcattgatgccttttggtttaaga 1740
|||||
Db 1681 agtttaatttttagtagaataagaatacacaagggcattgatgccttttggtttaaga 1740

QY 1741 tcggaatgtacaatatgtgcaatgagcagatatggagaaaaccttttgcgqgaaaaacat 1800

| | | | | | |
|----------|--|--|------|--|--|
| Db | 722 | gcgcgttcacgcccggagttccctggaaacaaagccctccaggttgacgacagacagatagtgc | 781 | | |
| QY | 851 | aaaatctgtgggcgagaaacgagatgaaagaaggaggagccattgtgacggtgaggggag | 910 | | |
| Db | 782 | aaaaccctaaagagcgagaccgagagctgaaagaaggaggaccattgtgacagtgagggaag | 841 | | |
| QY | 911 | gcctcagaatcttgagcccaagatggaacgagaggtgcccgaagaaagaggaatacgtg | 970 | | |
| Db | 842 | gcctcagaatcttgagcccaagatggaagaagagacgtgcccgaagaaagaggaatacgtg | 901 | | |
| QY | 971 | aagatcaatatgaataacatgaacagagatggaagcgctgycgagggggaagcagagcgggg | 1030 | | |
| Db | 902 | aagatgaatatgaataacgtatgaagaggaatagaaggcgtgycaggggggaagcagagcggg | 961 | | |
| QY | 1031 | ggaatggtatgaaagaagacgatctgcacccgcatgtgttaaaaaagaacattggtgaaaca | 1090 | | |
| Db | 962 | ggaatggtatgaaagaagacgatctgcacccgcaagtgctlaaaaaagaacattggtagaaca | 1021 | | |
| QY | 1091 | gacccctcacatctaacgaatccacagcgtgttcaactcaaaaactgcc-acgatctcaac | 1149 | | |
| Db | 1022 | gacccctgacatctacaacccctca-acgtgtgtcaactcaaaaactgccaacgactcaac | 1080 | | |
| QY | 1150 | cttctaactccttagtggtcttgacctagtgtcgtgaatalatggaatctctacaggaatgca | 1209 | | |
| Db | 1081 | cttctaactccttagtggtcttgacctagtgtcgtgaatalatggaatctctacaggaatgca | 1140 | | |
| QY | 1210 | ttgtttgtccctcaactacaacccaacgacacagcatcatatatgtcatgtgagggacgg | 1269 | | |
| Db | 1141 | ttgtttgtcgtctcaactacaacccaacgacacagcatcatatatcgtattgaggggacgg | 1200 | | |
| QY | 1270 | gctcacgtgcaagtgtgtgacacgcaacgccaacagagtgtacgcagagaggttcaagag | 1329 | | |
| Db | 1201 | gctcacgtgcaagtcgtgtgacacgcaacgccaacagagtgtacgcagagaggttcaagag | 1260 | | |
| QY | 1330 | ggtcacgttctgtgtgtgccacagaaactcgccgtgtgtggaagtcccagagcgagaac | 1389 | | |
| Db | 1261 | ggtcacgtgtgtgtgtgccacagaaactcgccgtgtgtggaagtcccagagcgagaac | 1320 | | |
| QY | 1390 | ttcgaatacgttgatccaagaacagattcaaggcccagcatagccaacttgcggtgaa | 1449 | | |
| Db | 1321 | ttcgaatacgttgatccaagaacagattcaaggcccagcatagccaacttgcggtgaa | 1380 | | |
| QY | 1450 | aactcctcatagataaactgcccggagaggtgtgtgcaaatcatatgtgcctcccaag | 1509 | | |
| Db | 1381 | aactcctcatagataaactgcccggagaggtgtgtgcaaatcatatgtgcctcccaag | 1440 | | |
| QY | 1510 | gagcagggcaagcgagcttaagaacaacaaccccttcaagttctcgttccaccttccag | 1569 | | |
| Db | 1441 | gagcagggcaagcgagcttaagaacaacaaccccttcaagttctcgttccaccttccag | 1500 | | |
| QY | 1570 | cagtcctcgaggggtgtgtgcttaa | 1593 | | |
| Db | 1501 | cagtcctcgaggggtgtgtgcttaa | 1524 | | |
| RESULT 4 | | | | | |
| ID | AAS08540 | AAS08540 standard; cDNA; 1524 BP. | | | |
| AC | AAS08540; | | | | |
| DT | 23-OCT-2001 | (first entry) | | | |
| DE | DNA encoding anaphylactic antigen Ara h 3. | | | | |
| KW | Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic; | | | | |
| XX | allergy; mast cell; basophil; mouse; ss. | | | | |
| OS | Mus sp. | | | | |
| FH | key | Location/Qualifiers | | | |
| FT | CDS | 1..1533 | | | |
| FT | /*tag= a | | | | |

```

FT      /product= "peptide antigen Ara h3"
FT      /note= "Coding sequence of amino acids 1-3 not given"
XX
XX
PN      WO200140264-A2.
XX
XX
PD      07-JUN-2001.
XX
XX
PF      06-DEC-2000; 2000WO-US33124.
XX
XX
PR      06-DEC-1999; 99US-0455294.
PR      23-JUN-2000; 2000US-0213765.
PR      27-SEP-2000; 2000US-0235797.
XX
XX
PA      (PANA-) PANACEA PHARM LLC.
PA      (UYAR-) UNIV ARKANSAS.
PA      (MOUN ) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX
XX
PI      Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
DR
XX
DR      WPI; 2001-381378/40.
XX
DR      P-PSDB; AAU04708.
XX
XX
PT      Antigenic fragments useful for reducing anaphylactic risk and reducing
PT      the severity and/or number of allergic symptoms in individuals
PT      sensitive to antigens, have reduced ability to bind Immunoglobulin E -
XX
XX
PS      Disclosure; Fig 11; 100pp; English.
XX
XX
CC      The sequence represents the coding sequence of anaphylactic antigen
CC      Ara h 3. Ara h 3 is an anaphylactic antigen (A), which was used to design
CC      antigenic peptides having a reduced ability to bind immunoglobulin E
CC      (IGE) as compared with the intact (A), or having a sequence substantially
CC      identical to a portion of sequence of an antigen that includes at least
CC      one IGE binding site, where at least one IGE binding site of the peptide
CC      is altered. The antigenic peptides are used in a composition which is
CC      useful for reducing risk or severity of allergic reaction to an antigen.
CC      This is done by identifying an individual at risk of allergic reaction to
CC      an antigen by identifying prior display of allergic symptoms when exposed
CC      to the antigen, or a familial relationship with an individual who
CC      previously displayed allergic symptoms when exposed to the antigen.
CC      Following this an antigen-specific IGE present on one or more mast cells
CC      or basophils in the individual's serum is identified. The individual is
CC      then contacted with a peptide corresponding to a portion of the
CC      antigen, which is selected, formulated, and delivered so that binding of
CC      the peptide to antigen-specific IGE is reduced as compared with IGE
CC      binding of intact antigen. The composition is also useful for treating
CC      and preventing allergic reactions.
XX
XX
SQ      Sequence 1524 BP; 455 A; 392 C; 396 G; 281 T; 0 other;
XX
XX
Query Match          74.9%; Score 1388; DB 22; Length 1524;
Best Local Similarity 95.3%; Pred. No. 0;
Matches 1452; Conservative 0; Mismatches 70; Indels 2; Gaps 2;
XX
QY      71  ggcagcagcgcggaggaatgctgtgccagtccagcgccctcaatgcgcaagagacctgaca 130
DB      2  ggcagcaaccggaaggagaacgctgtgccagttccagcgccctcaatgcgcaagagacctgaca 61
QY      131 accgcattgaatcggaaggcggttacattgagacttggaacccaacaacaggagttcg 190
DB      62  atcgcatatgaatcagaaggcggttacattgagacttggaacccaacaacaggagttcg 121
QY      191 aatgcgcgcgcgtccctctctcgtctagttctctccgcgcgaacgcccctctgtagccct 250
DB      122 aatgcgcgcgcgtccctctctcgtctagttctctccgcgcgaacgcccctctgtagccct 181
QY      251 totactccaatgtctcccaaggagatcttcatcaccagcaaggaaggagatactttggttga 310
DB      182 totactccaatgtctcccaaggagatcttcatcaccagcaaggaaggagatactttggttga 241
QY      311 tattccctgtgttctcttagcacatatgaagagcctgcacaacaaggaagcgcgatatcagt 370
DB      182 tattccctgtgttctcttagcacatatgaagagcctgcacaacaaggaagcgcgatatcagt 370

```


Db 242 tatlccctgtgtctccttagacactatgaaagcctcacacacaaggtcgtcgatctcagt 301
QY 371 cccaagaaccaccaagacgttttgcaagaagaagaccaaagccaacagcaacagatagtc 430
Db 302 cccaagaaccaccaagcgtctccaaaggagaagaaccaaaagccaacagcaacagatagtc 361
QY 431 accagaagtgcaaccglttccaatgaggggtgatctcatgtgcagttcccacgggtgtgctt 490
Db 362 accagaagtgcaaccglttccgatgaggggtgatctcatgtgcagttcccacgggtgtgctt 421
QY 491 tctggtcttaacaacgacacagcacatgatgtgtgtgtgttctcttactagcaccaaca 550
Db 422 tctggtcttaacaacgacacagcacatgatgtgtgtgtgttctcttactagcaccaaca 481
QY 551 acaacgacaaccagcttgatcagttccccagagatctcaatttgctgggaaaccagagc 610
Db 482 acaacgacaaccagcttgatcagttccccagagatctcaatttgctgggaaaccagagc 541
QY 611 aagagttcttaaggtaccagcaacaagcagacaaagcagacgaagaagcttaacatata 670
Db 542 aagagttcttaaggtaccagcaacaagcagacaaagcagacgaagaagcttaacatata 601
QY 671 gcccatacagcccgcatagtcggtccttagacgagaagcgltgaattcgcctcgagagc 730
Db 602 gcccatacagcccgcaaaagtcagccttagacaagaagagcgtgaatttagcctcgagagc 661
QY 731 agcacagccgcagagaagaacgagcaggaacaagaagaagacgaagtggaacaatcttca 790
Db 662 agcacagccgcagagaagaacgagcaggaacaagaagaagacgaagtggaacaatcttca 721
QY 791 gcggtctcacgcgcgaggttctctggaacaagccttcaggttgacgacagacagattgtgc 850
Db 722 gcggtctcacgcgcgaggttctctggaacaagccttcaggttgacgacagacagattgtgc 781
QY 851 aaaatctgtggygcgaagacgagagtgaaagaagggagccattgtgacggtgagggag 910
Db 782 aaaactaagagcgagacgagagtgaaagaagggagccattgtgacagtgagggag 841
QY 911 gcctcagaatcttgagcccaagatggaacgagaggtgcgcgaagaagaaggaatacagatg 970
Db 842 gcctcagaatcttgagcccaagatggaagaagagacgtgccgaagaagaaggaatacagatg 901
QY 971 aagatcaatatgataccatgaacagagatggaagcgctgycaggggaaagcagagcgggg 1030
Db 902 aagatgaatatgataccatgaacagagatggaagcgctgycaggggaaagcagagcgggg 961
QY 1031 ggaatggtatggaagagacgatctgcaccgcgatgtgttaaaaagaacattggtggaaca 1090
Db 962 ggaatggtatggaagagacgatctgcaccgcgatgtgttaaaaagaacattggtggaaca 1021
QY 1091 gatccctcacatctcagatctctcagcgctggttcaactcaaaactgcc-acgatctcaac 1149
Db 1022 gatccctgcacatctcaaccctca-acgtggttcaactcaaaactgccaacagatctcaac 1080
QY 1150 ctctcaatccttagtggcttggaactagtgctgaatatggaatctctacaggaatgca 1209
Db 1081 ctctcaatccttagtggcttggaactagtgctgaatatggaatctctacaggaatgca 1140
QY 1210 ttgtttgtccctcactacaacccaacgacacacagcatcatatgcatltgaggggagcgg 1269
Db 1141 ttgtttgtcgtcactacaacccaacgacacacagcatcatatgcatltgaggggagcgg 1200
QY 1270 gctcacgtgcaagtggtgacagcaacgccaacagagtggtacgacgagagcttcaagag 1329
Db 1201 gctcacgtgcaagtggtgacagcaacgccaacagagtggtacgacgagagcttcaagag 1260
QY 1330 ggtcacgttctgtgtgtgccaacagaacttcgcgctggtggaagtcgccagagcgagaac 1389
Db 1261 ggtcacgttctgtgtgtgccaacagaacttcgcgctggtggaagtcgccagagcgagaac 1320
QY 1390 ttcgataacgtgcatctcaagacagatctcaaggccagcatagccaacttgcgcggtgaa 1449
Db 1321 ttcgataacgtgcatctcaagacagatctcaaggccagcatagccaacttgcgcggtgaa 1380

QY 1450 aactcctcatagataacctgcgcgagaggtgtgcaaatctatatgctcccaagg 1509
Db 1381 aactcctcatagataacctgcgcgagaggtgtgcaaatctatatgctcccaagg 1440
QY 1510 gagcagggcaaggcagcttaagaacacacacccttcaagttcttggttccacttcaag 1569
Db 1441 gagcagggcaaggcagcttaagaacacacacccttcaagttcttggttccacttcaag 1500
QY 1570 cagttccgagggctgtggttaa 1593
Db 1501 cagttccgagggctgtggttaa 1524

RESULT 5
AAV17567
ID AAV17567 standard; cDNA; 1446 BP.
XX
AC AAV17567;
XX
DT 10-JUN-1998 (first entry)
XX
DE Coding sequence for Gy3 subunit of glycinin.
XX
KW Beta-conglycinin; soybean seed protein; transgenic plant; glycinin;
KW seed storage protein profile; ss.
XX
OS Glycine max.
XX
PN WO9747731-A2.
XX
PD 18-DEC-1997.
XX
PF 10-JUN-1997; 97WO-US09743.
XX
PR 14-JUN-1996; 96US-0019940.
XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Fader GM, Kinney AJ;
XX
DR WPI; 1998-052298/05.
XX
PT Suppression of specific classes of soybean seed protein genes -
XX
PS useful to change seed storage protein profiles of transgenic plants
XX
XX Disclosure; Page 36-37; 58pp; English.

CC This sequence represents the coding sequence for the Gy3 subunit of the
CC soybean seed protein, glycinin. The method of the invention is for
CC reducing the quantity of a soybean seed storage protein (A), such as
CC beta-conglycinin, in soybeans. It comprises: (a) constructing a chimeric
CC gene comprising: (i) a nucleic acid fragment encoding a promoter that is
CC functional in the cells of soybean seeds; (ii) a nucleic acid fragment
CC encoding all or a portion of (A) placed in sense or antisense orientation
CC relative to the promoter of (i); and (iii) a transcriptional termination
CC region; (b) creating a transgenic soybean cell by introducing into a
CC soybean cell the chimeric gene of (a); and (c) growing the transgenic
CC soybean cells of (b) under conditions that result in expression of the
CC chimeric gene of (a); where the quantity of one or more members of a
CC class of (A) subunits is reduced when compared to soybeans not containing
CC the chimeric gene of (a). The method is used to construct transgenic
CC soybean lines where the expression of genes encoding (A) are modulated to
CC effect a change in seed storage protein profile of transgenic plants.
CC Modification of the seed storage protein profile can result in the
CC production of novel soy protein products with unique and valuable
CC functional characteristics.
XX
SQ

Sequence 1446 BP; 426 A; 361 C; 354 G; 305 T; 0 other;

Query Match 28.5%; Score 527.2; DB 19; Length 1446;
Best Local Similarity 62.2%; Pred. No. 6.7e-119;

[illegible]

| | | | |
|----|------|--|------|
| QY | 1093 | tcocctcacatctaagatcctcagcgcgtggttcaactcaaacctgcacgatctcaacct | 1152 |
| | | | |
| Db | 946 | tcacctgacatctctcaaccctcaagctgtgtagcatcacacacgcgtcacagcctgcacttc | 1005 |
| QY | 1153 | ctaactcctaagtggtgcttggacttagtgcgtgaatatgynaatctctacagaagatgcatly | 1212 |
| | | | |
| Db | 1006 | ccagccctctcgtggtccaactcagtcagtcgccagtttgatcacctccgacaagaatgctatg | 1065 |
| QY | 1213 | tttgtccctcacatacaacccaacgcacacagcatcatatatgcatlttgagggagcggt | 1272 |
| | | | |
| Db | 1066 | ttcgtgcacactacaacctgaaacgcacaacataatacgcattgaaatgyacgycga | 1125 |
| QY | 1273 | cacgtgcagtggtgagcacgacaacgycacaacagagltgacgacgagaggtccaagaggt | 1332 |
| | | | |
| Db | 1126 | ttgttacaagtggtgaaatgcatggtgagagagtglttgatgagagctgcaagagga | 1185 |
| QY | 1333 | cacgttcttgtgtgcccacagaacttcgcgctggtgaggaagtcaccagcgagaacttc | 1392 |
| | | | |
| Db | 1186 | cagtggttaattgtgcccacaaaacttgcggtggtgcgaagatcacagagcgacaacttc | 1245 |
| QY | 1393 | gaatacgtgacattcaagacagatccaagcccagcatagccaacttgcggtgaaac | 1452 |
| | | | |
| Db | 1246 | gagtatgttcattcaagaaccaatgatagaacctcgatcggaaccttgcaagtgcaaac | 1305 |
| QY | 1453 | tccttcataagataacctgcggagaggtgtgcaaatcatatgagcctcccaaggag | 1512 |
| | | | |
| Db | 1306 | tcattgttgaacgcattgcggaggaagtgatcagcaaacctttaacctaaaggaagcag | 1365 |
| QY | 1513 | caggcaaggcagcttaagaacaacaacccctcaagttcttcgttccaccttccagcag | 1572 |
| | | | |
| Db | 1366 | caggccaggcaggtccaagaacaacaaccccttcaagcttcctggttccacctaaagagtc | 1425 |
| QY | 1573 | tctccgaggcgtgtgctta | 1592 |
| | | | |
| Db | 1426 | cagaggaagtggtgctta | 1445 |

```

RESULT      6
AAZ92638
ID    AAZ92638 standard; cDNA; 1743 BP.
XX
XX      AAZ92638;
AC
XX      05-JUN-2000 (first entry)
DT
XX      cDNA encoding soybean glycinin Alab1b.
DE
XX      Soybean; glycinin; transgenic rice; glutelin promoter; seed expression; ss.
KW      processing; nutritional quality; antilipemic; cholesterol;
KW      ss.
XX
OS      Glycine max.
XX
XX      Key      Location/Qualifiers
FH      CDS      52..1539
FT      /*tag= a
FT      /product= "Soybean glycinin Alab1b"
XX
XX      WO200008161-A1.
XX
XX      17-FEB-2000.
PD
XX
XX      04-MAR-1999; 99WO-JP01057.
PF
XX
XX      07-AUG-1998; 98JP-0223897.
PR
XX
XX      (NORO ) JAPAN MIN AGRIC FORESTRY & FISHERIES.
PA      (BIOO-) BIO-ORIENTED TECHNOLOGY RES ADVANCEMENT.
XX
XX      Takaiwa F, Utsumi S, Katsube T;
PI
XX
XX      WPI; 2000-205714/18.
DR      P-PSDB; AAY80994.

```


Db 1322 tgatcggcactctcttgcagggcaaacatcatgttgaacgcattaccagagagtgtaltc 1381
QY 1487 caatlcataatgacctcccaaggagcaggcaagcagcttaagaacaacaacccctca 1546
Db 1382 agcacacttccaacctaaagaagcagcaggccagcagataaagaacaacaaccccttca 1441
QY 1547 agttcttcgtccacctttccagcagtcctccgaggctgtgctta 1592
Db 1442 agttcctgttccacctcagcagagtcctcagaagagagctgtgctta 1487
RESULT 8
AAQ05359
ID AAQ05359 standard; DNA; 1712 BP.
XX
AC AAQ05359;
XX
DT 04-DEC-1990 (first entry)
XX
DE Glycinin subunit precursor A2B1a.
XX
KW Glycinin; glycine; pI6Vneo1103; ds.
XX
OS Glycine hispida.
XX
PN JP02156889-A.
XX
PD 15-JUN-1990.
XX
PF 08-DEC-1988; 88JP-0310553.
XX
PR 08-DEC-1988; 88JP-0310553.
XX
PA (NORO) NORINSO.
XX PA (KIRI) KIRIN BREWERY KK.
DR WPI; 1990-228488/30.
XX
XX Recombinant plasmid - obt'd. by recombining glycinin gene to
PT plasmid for plant.
XX
PS Claim 6; Page 561; 20pp; Japanese.
XX
CC Glycinin precursor gene may be used with exogenous promoter and
CC terminator in plasmid pI6Vneo1103 to transform plants, improving
CC nutritional value.
XX
SQ Sequence 1712 BP; 522 A; 400 C; 414 G; 375 T; 1 other;
Query Match 27.6%; Score 510.6; DB 11; Length 1712;
Best Local Similarity 61.7%; Pred. No. 8.1e-115;
Matches 977; Conservative 0; Mismatches 474; Indels 132; Gaps 5;
QY 17 agcttccttttggcttttggcttctctagttctctggagctagcagcatctcctcaggcagc 76
Db 52 agctgttcttctcccttggtttcccttcttccag--tggctgtctgccttgagagagc 108
QY 77 agccggagagaaatgctgcagcttccagcgctcaatgcgcagagacctgacaacccgca 136
Db 109 aggcacagcaaaatgagtgccagatccaaaagctgaatgcctcaaaccaaggtaacctga 168
QY 137 ttgaatcggaggggcggttacatgagacttggaacccaacaacaggagtgcaatgcg 196
Db 169 tagagtcggaagtggttcatgtagaccttggaacccaacaacaagccattcagtggt 228
QY 197 ccggcgtcgacctctctcgtctagtcctccgcgcgcaacgacctgtagaccttctact 256
Db 229 ccggtgtgcccctctctcgtcgcaccccttaaccgcgcaatgccttcgtagaccttcccta 288
QY 257 ccaatgctccccagagatcttcatccagcaaggaaggatacttgggtgataattcc 316
||||| ||||||||| ||||| ||| ||||| | ||| |||||||||

Db 289 ccaacggtccccaaggaaatctacatacaacaaggtaatgtatttttggcatgatattcc 348
QY 317 ctggtgtccttagacatatgaagagcctgcacaacaaggagccgatatacgttcccaa 376
Db 349 cgggtgtccttagcacttatcaagagccgcaagaa----- 383
QY 377 gaccacaagaacgtttgcaagaagaagccaagccaacagcaacaagatagtcaccaga 436
Db 384 -----tctcagcaacgaggaagcgaagccagaggccccaagaccgttcaccaa 429
QY 437 agtgtcacogtttcaatgaggtgatctcatattgcagttcccaacgggtgtgtcttctgc 496
Db 430 aggtacatcgtctcagagaggtgtgatttgatcgcagtgccctactgtgtgtcagtgtgga 489
QY 497 tgtacaacgaccaacgacactgatatgttgtcgtgttctcttactgacaccaacaacag 556
Db 490 tgtacaacaatgaagacactcctgtgtgtgcggttcttattatlgacaccaacagcttg 549
QY 557 acaacacgcttgatcagttccccaaggagatttcaatttggctggaaccaacgagcaagagt 616
Db 550 agaaccagctcgaccagatgcttagagattctatcttctgtggaaccaagagcaagagt 609
QY 617 tcttaaggtaccagcaacaagcagacaagaagcagacgaagaagcttaccatatagccat 676
Db 610 ttcctaaatatcagcagca----- 628
QY 677 acagcccgcatagtcgcttagacgagaagagcgtgaatttcgccctcgaggagcagcaca 736
Db 629 -----gcagcaaggaggttccc 645
QY 737 gccgcagagaaacgagcagagacaagaagaagaaagcgaaggttgaaacatcttcagcgct 796
Db 646 aaagccagaaagaaagcaacaagaagaagaagaagaagcaacatatgtagtgct 705
QY 797 tcaagccggaagttcctggaacaagccttccaggttgacgacagacagatgtgcaaatc 856
Db 706 tcgcccctgaattcttgaagaagcgttc--g'gcgtgaacatgacagatgtgagaacc 762
QY 857 tgtggygcgagaacgagagtgaaagaagagggagccattgtgacggtgagggagcgctca 916
Db 763 tacaaggtgagaacgagagagagatagtgagcagcatgtgacagtgaaaggaggtctaa 822
QY 917 gaatcttgagccagatggaacgagaggtgcccgcgaagaagaagatacgtga---ag 973
Db 823 gagtcacagctccagccatgaggaagccacgcgaagaagaagaatgatgatgaggrag 882
QY 974 atcaatatgaataccatgaacagagatgtgaagggcgtgagcaggggaagcagggcgga 1033
Db 883 agcagccacagtgctgtgagacagacaagaaggttgccaacgccaagaagcaaaagagcagaa 942
QY 1034 atggtattgaagaaacgatatctgcacccgcgatgtgttaaaaagaacattggtgaaacagat 1093
Db 943 atgcatgtatgagaaacatttgcaacaatgagacttcgccaanaacattgylcagaattcat 1002
QY 1094 cccctcacatctacgatacctcagcgcgtgttcaactcaaaactgccacgatctcaaccttc 1153
Db 1003 cactgacatctacaacccctcaagctgtgtagcatcacaaacgcgcaaccagccttgacttcc 1062
QY 1154 taatccttagtggcttggacttgaatgtgaatatggaatctctacaggaatgcatgt 1213
Db 1063 cagccctctgtcttctcaaaactcagctgcagctcagatgtagtaccctcgcaagaatgtcagt 1122
QY 1214 ttgtccctcactacaacaacaacgacacacagcatcatatatgtcatgtgagggagcgctc 1273
Db 1123 tcgtgccacactacaacccctgaacgcggaacagcatataatagcatgtgaatggcgcgcat 1182
QY 1274 acgtgcaagtgtgtgacagcaacgcggaacagagtgtagcagcagaggagcttcaagggttc 1333
Db 1183 tggtaacaagtgtgaattgcaatgtgtgagagagtggtttagtggagagcgtcgaaggggag 1242
QY 1334 acgttcttggtgtgccacagaacttcgcgctgtgctgggaagtcgccagagcgagaacttcg 1393
Db 1243 ggtgtgtgactgttccacaanaaacttgcggtgtgctgcataaatccagagcgataacttg 1302

| | | | |
|----|------|--|------|
| QY | 1394 | aatcactgtgcattcaagaacagattccaaggccaccagcatagccaactttgccggtgtgaaaaact | 1453 |
| | | | |
| Db | 1303 | agtatgtgtcattccaagaaccaatgatatagacccttcgatcgaaaccttgcagggtgcaaac | 1362 |
| QY | 1454 | ccttcatatagataaacctggccggaggaagtgtgtgcaaattcatatatgacctccaaggagc | 1513 |
| | | | |
| Db | 1363 | cattgttgaaacgcattcggccagaggaagtgtattcagcacactttaacctaaagayccagc | 1422 |
| QY | 1514 | aggccaagggcagcttaagaacaacaaccccccttcaagrttcttgttccacccttttcagcagt | 1573 |
| | | | |
| Db | 1423 | aggccaagggcaggtgaagaacaacaacacccttccagcttcccttgttccaccctcaggagcttc | 1482 |
| QY | 1574 | ctccgagggtgtgtgcttaaaa | 1596 |
| | | | |
| Db | 1483 | agaggagaagctgtgtgcttagaga | 1505 |

| | |
|----------|---|
| RESULT | 9 |
| AAV17566 | |
| ID | AAV17566 standard; cDNA; 1458 BP. |
| XX | |
| AC | AAV17566; |
| XX | |
| DT | 10-JUN-1998 (first entry) |
| XX | |
| DE | Coding sequence for Gy2 subunit of glycinin. |
| XX | |
| KW | Beta-conglycinin; soybean seed protein; transgenic plant; glycinin; |
| KW | seed storage protein profile; ss. |
| XX | |
| OS | Glycine max. |
| XX | |
| PN | WO9747731-A2. |
| XX | |
| PD | 18-DEC-1997. |
| XX | |
| PF | 10-JUN-1997; 97WO-US09743. |
| XX | |
| PR | 14-JUN-1996; 96US-0019940. |
| XX | |
| PA | (DUPO) DU PONT DE NEMOURS & CO E I. |
| XX | |
| PI | Fader GM, Kinney AJ; |
| XX | |
| DR | WPI; 1998-052298/05. |
| XX | |
| PT | Suppression of specific classes of soybean seed protein genes - |
| PT | useful to change seed storage protein profiles of transgenic plants |
| XX | |
| PS | Disclosure; Page 35-36; 58pp; English. |
| XX | |
| CC | This sequence represents the coding sequence for the Gy2 subunit of the |
| CC | soybean seed protein, glycinin. The method of the invention is for |
| CC | reducing the quantity of a soybean seed storage protein (A), such as |
| CC | beta-conglycinin, in soybeans. It comprises: (a) constructing a chimeric |
| CC | gene comprising: (i) a nucleic acid fragment encoding a promoter that is |
| CC | functional in the cells of soybean seeds; (ii) a nucleic acid fragment |
| CC | encoding all or a portion of (A) placed in sense or antisense orientation |
| CC | relative to the promoter of (i); and (iii) a transcriptional termination |
| CC | region; (b) creating a transgenic soybean cell by introducing into a |
| CC | soybean cell the chimeric gene of (a); and (c) growing the transgenic |
| CC | soybean cells of (b) under conditions that result in expression of the |
| CC | chimeric gene of (a); where the quantity of one or more members of a |
| CC | class of (A) subunits is reduced when compared to soybeans not containing |
| CC | the chimeric gene of (a). The method is used to construct transgenic |
| CC | soybean lines where the expression of genes encoding (A) are modulated to |
| CC | affect a change in seed storage protein profile of transgenic plants. |
| CC | Modification of the seed storage protein profile can result in the |
| CC | production of novel soy protein products with unique and valuable |
| CC | functional characteristics. |
| XX | |
| SO | Sequence 1458 BP; 436 A; 348 C; 367 G; 307 T; 0 other; |

[illegible]

| | | | |
|----|------|---|------|
| QY | 1034 | atgttatggaagagacgacatctgcaccgcgcatgtgtttaaaaaagaacatttgttgaacacagat | 1093 |
| | | | |
| Db | 899 | atgacatgatgagaccatttgcacaatagagacttcgcacaaacatttgttcagaatccat | 958 |
| QY | 1094 | ccccctacatctacgataccctcagcgcgtgttcaactcaaaactgccaagatctcaaccttc | 1153 |
| | | | |
| Db | 959 | cacctgacatctacacaccctcaagctgtgtagcatcacacaccgcaccacgcttgacttcc | 1018 |
| QY | 1154 | taatccttagtgtgtcttgcacttagtgtctgaatatgtgaatactctacaggaatgcattgt | 1213 |
| | | | |
| Db | 1019 | cagccctctgtgccttctcaaaactcagtcagtcgccagtatgtacactccgcaagaatgtctatgt | 1078 |
| QY | 1214 | ttgtccctcaactacaacaaccacacgcacacagcatcatatgtcaattgagggagcggctc | 1273 |
| | | | |
| Db | 1079 | tcgtgccacactacacccctgaacgcgacagcatatatacgcattgtaatgttgcgggcacat | 1138 |
| QY | 1274 | acgtgcgaagtgtgtgacacgcaacgcgcaacagagtgtacgacgagagcttcaagaggctc | 1333 |
| | | | |
| Db | 1139 | tgttacaaagtgtgtgaatttgcaatgtgtgagagagtgtttgtgagagagctgcgaaggaggag | 1198 |
| QY | 1334 | acgttcttgtgtgtgccacagaaacttcgccgtgtgtcgggaagtcgccagagcgaagaactcg | 1393 |
| | | | |
| Db | 1199 | gggtgtcgtatcgttccacaaaactttgcgtgtgtgtcaaaatccagagcgaatbaacttgg | 1258 |
| QY | 1394 | aatacgtgtgcattcaagaacagatttcaaggccacagcatagccaactttgcggtgtgaaaact | 1453 |
| | | | |
| Db | 1259 | agtatgttcatlcaagaaccaatgatagaccctcgcgtcgaacaccttgcaggggccaact | 1318 |
| QY | 1454 | ccttcatagataaaccttgcgcgagagagtggttgcacaattcatatgtgcctcccaagggagc | 1513 |
| | | | |
| Db | 1319 | catgttgtgaacgcatgtgccagaggaagtgtatccagcacacttttaaccctaagaagccagc | 1378 |
| QY | 1514 | aggccaagcgacgttaagaacaacaacaccccttcaagtcttctgttccaccttlttcagcagt | 1573 |
| | | | |
| Db | 1379 | aggccaagcgacgttgaagaacaacaacaccccttccagcttccctgttccacctcaggaagtc | 1438 |
| QY | 1574 | ctccgagagcctgtgtgcctta | 1592 |
| | | | |
| Db | 1439 | agaagagaagctgtgtgcctta | 1457 |

RESULT 10
 AAQ05360
 ID AAQ05360 standard; DNA; 1746 BP.
 AC AAQ05360;
 DT 04-DEC-1990 (first entry)
 DE Glycyltln subunit precursor AlaB1b.
 KW Glycyltln; glycine; pLGvne01103; ds.
 OS Glycine hispida.
 PN JP02156889-A.
 PD 15-JUN-1990.
 PF 08-DEC-1988; 88JP-0310553.
 PR 08-DEC-1988; 88JP-0310553.
 PA (NORQ) NORINSHO.
 PA (KIRI) KIRIN BREWERY KK.
 DR WPI; 1990-228488/30.
 PT Recombinant plasmid - obtd. by recombining glycyltln gene to
 PT plasmid for plant.
 XX
 Claim 7; Page 561; 20pp; Japanese.

| | |
|----|---|
| XX | Glycinin precursor gene may be used with exogenous promoter and terminator in plasmid pGlvneol103 to transform plants, improving nutritional value. |
| CC | |
| CC | |
| CC | |
| XX | |
| SQ | Sequence 1746 BP; 553 A; 421 C; 396 G; 374 T; 1 other; |

| | | | | |
|---------------------------|--------|---------------------|-------------|--------------|
| Query Match | 27.1%; | Score 501.6; | DB 11; | Length 1746; |
| Best Local Similarity | 61.7%; | Pred. No. 1.3e-112; | | |
| Matches 992; Conservative | 1; | Mismatches 459; | Indels 156; | Gaps 6; |

| | | | |
|----|-----|---|-----|
| QY | 17 | agcttctctttgtctttgtcttcttaagttctgtggaagctagcagcatctcctt-----ca | 70 |
| | | | |
| Db | 62 | agctagtttttcccttgttttctgtcttcttcagtggtctgtctgtcgtcttccagttcca | 121 |
| QY | 71 | ggcagcagccggaagagaatgcgtgcccagttccagcgcgcctcaatgctgcagagacctgaca | 130 |
| | | | |
| Db | 122 | gagaagcgcctcagcaaaaacgagtgccagatccaaaaactcaatgctccctcaaacggyta | 181 |
| QY | 131 | accgcattgaatcggagggcggtttacattgagacttggaaacccaacacagaggttcg | 190 |
| | | | |
| Db | 182 | accgtatagatcagaagaggaggtcatcttagagacatlygaaccctaacaacaagccattcc | 241 |
| QY | 191 | aatgcgcgcgcgtcgcgcctctctcgttagtgcctccgcgcgaacgcgcctcgttagccct | 250 |
| | | | |
| Db | 242 | agtgccgcggtgtgtgcctctctccctgcaccctcaaccgcgaacgcgcctcgttagaccct | 301 |
| QY | 251 | tctactccaatgctccccaaggagatcttcatccagcaaggaagggtatacttgggttga | 310 |
| | | | |
| Db | 302 | cctacaccaaagcgtccccaagaataatacatccaacaagtaagggtatltttcgcattga | 361 |
| QY | 311 | tattccctgtgtgtccttagcacatatgaaagacgttcgacacaacaaggaacgcgatatcagt | 370 |
| | | | |
| Db | 362 | tatacccggtgtgttcttagcacatttgaagagcctcaacaac----- | 403 |
| QY | 371 | cccaaaagcaccacaagaagcgtttgcaagaagaagaccacaagccacaacagataagtc | 430 |
| | | | |
| Db | 404 | -----ctcaacaaaaggagagaagcagcagaccacaagaacgcgtc | 442 |
| QY | 431 | accagaaggtgcaccgctttcaatgaggtgtatctcatctgcagttcccacgcgtgttgcct | 490 |
| | | | |
| Db | 443 | accagaagatctcataactccagagagagggtgattgtatcgcagtgccctactgtgtgcatt | 502 |
| QY | 491 | tctgctgttacaacgacccaacgacactgtatgttgtgtgttctcttacttgacaccaaca | 550 |
| | | | |
| Db | 503 | ggtgagatgtacaacaatgaaagacactcctgtgtgtgcgctttcattattgacaccaaca | 562 |
| QY | 551 | acaaagacaacacgcttgcatacagttccccaaggagattcaatttgcgtggygaaccaagac | 610 |
| | | | |
| Db | 563 | gcttgagaaacacgctcgcacacagatgctcctaggaattctatcttgytggaaccaagac | 622 |
| QY | 611 | aagagttcttaaggtaccagcaacaagaacagaacaagcagacgaagaaagccttaccatata | 670 |
| | | | |
| Db | 623 | aagagtttctaataatatacgaaga----- | 647 |
| QY | 671 | gcccatataagcccgcataagtcggccttagacgagaagagcgtgaatttcgcccctcgagac | 730 |
| | | | |
| Db | 648 | -----gcaagagagtc | 658 |
| QY | 731 | agcacagccgcagagaacgagcaggaacaagaagaagagcgaaggtgynaacatcttca | 790 |
| | | | |
| Db | 659 | atcaaaagccagaaaggaagacatcacgaagaagaagaagaacgaaggaagcagcatattga | 718 |
| QY | 791 | gcgcgtcaacgcgcgagttcctctggaacaagccttccaggttgcagacagacagattgtgc | 850 |
| | | | |
| Db | 719 | gtggtcttcaacccttgaatttcttgaacatgcatctcagcgt---ggaacaagcagatagaaga | 775 |
| QY | 851 | aaaaatctgtggtggtgcgagaacgagagtgaagaagagggagccattgtgacggtgaggygaag | 910 |
| | | | |
| Db | 776 | aaaaacctacaagagagagaacgaaaggggaagacaaagggagccaattgtgacagtgaaagac | 835 |
| QY | 911 | gcctcagaatcttgaagcca-----gatgaaacgaagaggtgcccagcaag | 955 |

DB 836 gtctcagcgtgtataaaaccaccacgcagcagcagcaacaagaaccccaaggaaggaag 895
QY 956 aagaggaatacagatgaagatcata-----tgaatacatgaacagagatggaagc 1006
DB 896 aagaagaagagatgagaagccacagtgcaagggtlaaagacaacactgcgaacgcccc 955
QY 1007 gtgcgaggggaagcagaagcgcggggaatgtatgaagagacgatctgcaccgcatgtg 1066
DB 956 gaggaagccaagaagaagaagcagaagaatgtgcattgacgagaccatatgcaccatgagac 1015
QY 1067 ttaaaagaacattgtggaacaacagatcccccatactacgatccctcagcgctgttca 1126
DB 1016 ttccgcaacaacattgcccagacttcataccttgacaatacaccctcaagccggtagcg 1075
QY 1127 ctcaaaactgcacagatcacaaccttaacttagtggtggttgacttagtctgaat 1186
DB 1076 tcacaacgcgcacacagccttgacttcaccagccctctcgtgctcagactcagtgtgat 1135
QY 1187 atggaatctctacagaatgcatgtgtgtccctcactacaacaacacagcacacagca 1246
DB 1136 ttgggtctctccgaagaatgcaatgtctgtgcccacactacaacctgaacgcgaacagca 1195
QY 1247 tcataatgcatgtgagggagcgggtcaccgtgcagatggtgagacagcagcgaacagag 1306
DB 1196 taataacgcatgtgaatgcagcggcattgatatacaagtgtgtaattgcacgcgtgagagag 1255
QY 1307 tgaacgacgagagcctcaagagggtcacggtctctgtgtggtgccacagaactccgcgtg 1366
DB 1256 tgttgatgagagagctgcagaaggagcggtgtgtgtcgtgcacaaacttggtgtg 1315
QY 1367 ctggggaagtcaccagagcgaagaacttcgaatacgtgcatccaagacagattcaaggccca 1426
DB 1316 ctgcaagatcacagaagtgaacaacttcgagtatgttcattcaagaccaatgatataccca 1375
QY 1427 gcataagccaacttgcgcggtgaaaactccttcataagataaactgcgcggaagagtggtt 1486
DB 1376 tgatcggcactcttgcaaggcacaactcattgttgaaacgcatlaacagaagagtgattc 1435
QY 1487 caaatcataltgccctcccaaggagcagcagcagcagcagcagcagcagcagcagcagc 1546
DB 1436 agcacacttcaaccataaaagcagcagcagcagcagcagcagcagcagcagcagcagc 1495
QY 1547 agtctctcgtccaccttttcagcagctctccgagcgctgtggttaaa 1594
DB 1496 agtctcgtgtccacactcagcaggtctcagaagaagagctgtggttaga 1543

RESULT 11
AAV17569
ID AAV17569 standard; cDNA; 1551 BP.
XX AAV17569;
AC AAV17569;
XX 10-JUN-1998 (first entry)
DT 10-JUN-1998 (first entry)
XX Coding sequence for Gys subunit of glycinin.
DE Beta-conglycinin; soybean seed protein; transgenic plant; glycinin;
XX seed storage protein profile; ss.
KM Glycine max.
OS Glycine max.
XX WO9747731-A2.
PN 18-DEC-1997.
PD 10-JUN-1997; 97WO-US09743.
PF 14-JUN-1996; 96US-0019940.
PR (DUPO) DU PONT DE NEMOURS & CO E I.
XX

PI Fader GM, Kinney AJ;
XX WPI; 1998-052298/05.
DR
XX Suppression of specific classes of soybean seed protein genes -
PT useful to change seed storage protein profiles of transgenic plants
XX
XX Disclosure; Page 38-39; 58pp; English.
PS
XX This sequence represents the coding sequence for the Gys subunit of the
XX soybean seed protein, glycinin. The method of the invention is for
CC reducing the quantity of a soybean seed storage protein (A), such as
CC beta-conglycinin, in soybeans. It comprises: (a) constructing a chimeric
CC gene comprising: (i) a nucleic acid fragment encoding a promoter that is
CC functional in the cells of soybean seeds; (ii) a nucleic acid fragment
CC encoding all or a portion of (A) placed in sense or antisense orientation
CC relative to the promoter of (i); and (iii) a transcriptional termination
CC region; (b) creating a transgenic soybean cell by introducing into a
CC soybean cell the chimeric gene of (a); and (c) growing the transgenic
CC soybean cells of (b) under conditions that result in expression of the
CC chimeric gene of (a); where the quantity of one or more members of a
CC class of (A) subunits is reduced when compared to soybeans not containing
CC the chimeric gene of (a). The method is used to construct transgenic
CC soybean lines where the expression of genes encoding (A) are modulated to
CC effect a change in seed storage protein profile of transgenic plants.
CC Modification of the seed storage protein profile can result in the
CC production of novel soy protein products with unique and valuable
CC functional characteristics.
XX
SQ Sequence 1551 BP; 449 A; 408 C; 376 G; 318 T; 0 other;

Query Match 10.9%; Score 201.8; DB 19; Length 1551;
Best Local Similarity 51.6%; Pred. No. 1.7e-39;
Matches 701; Conservative 0; Mismatches 597; Indels 61; Gaps 8;
QY 83 agagaatgcgtgccagttccagcgcctcaatgagcagagactgacaaccgcatgaaat 142
DB 86 agtcaacgagtgccaactcaacaactcaacgcgttggaaccgcgaaccgcggttgagt 145
QY 143 cgaagcggttaacattgagacttggaaccccaacaacagagattcgaatgcgcgcg 202
DB 146 ccgaagtggtctattgaacatggaacttcaacaccctgagctgcaatgcgcgcg 205
QY 203 tcgcctctctcgttagtctctccgcgcgaacgccccttgtagccttctactcaatg 262
DB 206 tcaactgttcccaacgcaccctcaaccgcgaacggtcccaactgcatctactactc 265
QY 263 ctcccaagagatcttcataccagcaaggaaggatatttggttgatatccctggt 322
DB 266 atcccaaatgatcaattgtcgttcaaggaaggagcaatgtgattgtcattccggat 325
QY 323 gtccatgacatatgaagagcctgcacacaagaagcgcgcatatcagttcccaagacc 382
DB 326 gtcccgagacgtttgagaagcacaacaatacaagcaga----- 366
QY 383 caagacgttgcagaagaagacccaagaagcacaagaagatagtcacagaaggtgc 442
DB 367 -----agaggtctaaggtcacagcagcaactacaagacagtcacagaagatc 415
QY 443 accgttcaatgaggtgatctcatgtgcagttcccaaccggtgtgtcttctgtgtaca 502
DB 416 gtcaacttcaatgaagagagcgtactagtatctctctgtgtcttcttactgacctata 475
QY 503 acgacacgacactgtgtgtgtgttcttcttactgacacaacaacaagaacaac 562
DB 476 acaactgagatgaaccaggtgtgtgcatacagtcctcttgacacctccaactcaacatc 535
QY 563 agcttgatcagttcccaaggagatccaatttgctggtggaaccagcagagagttcttaa 622
DB 536 agcttgatcaaaaccccgagtagtatttacccttgctggtggaaccagagatagagcatcccg 595
QY 623 ggtaccagcaacaagcagacaagaagcagagcagaagaagcttaccatatagcccatagcgc 682

Db 596 agaccatgcaacaa--cagcagcagcaagaagtcattgtgtgacgcgaagcaggggcaaca 653
QY 683 cgcatagtcgg-cctagaagaaagcgtgaatttcgcctcgagcagcacagccgc 741
Db 654 ccgacagcagggaagaaggtgctgctcagtgctcagcaaatcttctagc 713
QY 742 agagaacgagcaggaacaagaagaagcgaaggtgnaaatcttcag----cggtc 797
Db 714 acaatcttcaacaccacagagcacacagctgagaaacttcggtctccagatgacgaag 773
QY 798 cagccggagttccttgaaacaagccttcaggttgacgacagacagattgtgcaaatct 857
Db 774 gaagcagatcgtgacagtgaggaggccctcagcgttatcagcccccaagtggcaagaaca 833
QY 858 gtgggcgagaagcagagagtgaagaagaggagccattgtgacggtgagggagccctcag 917
Db 834 agaagcagaagacgaagacgaagcgaagaataatgacgcgcctcttatctccacg 893
QY 918 aatcttgagcccaagatggaacgaggtgc--gacgaagaaagagaatacagatgaaga 974
Db 894 acgaccaagccatggaagcatgaagatgacgagcagcgaagaaagaatacacc 953
QY 975 tcaatatgaataaccatgaacaagatggaag-----gcgtggcagggg 1016
Db 954 tcgtctgtatcacccctccacagcgaccaagcaggcccgacaacaagaaccacgtggaag 1013
QY 1017 aagcagaagcggggggaattgtattgaagagacgactcgcgcgcatgtgttaaaaaagaa 1076
Db 1014 aggatgtcagactagaatgggggttgagaaaaataattgcacatgaagcttcacgagaa 1073
QY 1077 cattgtgtgaaacagatcccctcacatctaagatcctcagcgtgtgttcaactcaaaactg 1136
Db 1074 cattgtctgccttcacgtgtgtacttctacaaccccaaaagctgtgtcgtattgacccct 1133
QY 1137 ccaagatctcaaccttctaactccttagtgtgtgactgtgactgtgctgaatatgaaatct 1196
Db 1134 caacagctcaccctccacgcccctccgccaattgcgactcagtgcccgaatatgtgtcct 1193
QY 1197 ctacaggaatgcatgtttgtlccctcactacaacaaccaacgcacacagcatcatatgtc 1256
Db 1194 ctacaggaatggaatttactctccagatttgaacttgaacgcgacacagtgtgacg-atga 1252
QY 1257 attgaaggagcgggtctacgctgcaagtgtgtgacagcaacgcgcaacagagtgtacgca 1316
Db 1253 ctgagagggaagaaga--gltagagtgtgtaactgccaagggaatgcagtgttcgacgg 1310
QY 1317 ggagcttcaagagggtcactgttctgtgtgtgcccacagaaacttcgcccgtgtggtggaagtc 1376
Db 1311 tgaactaaggagggaacaattgtctagtgtgtccgcgacgaacccccgcgtgtgtgacgaag 1370
QY 1377 ccagagcgagaacttcgaatacgtgcattcaagacaga 1415
Db 1371 gggagaacaagattggaatatgtagtgttcaagaacaca 1409

RESULT 12
AAQ05357
ID AAQ05357 standard; DNA; 1786 BP.
XX
AC AAQ05357;
XX
DT 04-DEC-1990 (first entry)
XX
DE Glycinin subunit precursor A3B4.
XX
KW Glycinin; glycine; PLGVneol103; ds.
OS Glycine hispida.
XX
PN JP02156889-A.
XX
15-JUN-1990.

XX 08-DEC-1988; 88JP-0310553.
PF
XX 08-DEC-1988; 88JP-0310553.
PR
XX (NORO) NORINSO.
PA (KIRI) KIRIN BREWERY KK.
XX WPI; 1990-228488/30.
DR
XX
PT Recombinant plasmid - obt'd. by recombining glycinin gene to
PT plasmid for plant.
PS Claim 4; Page 561; 20pp; Japanese.
XX
CC Glycinin precursor gene may be used with exogenous promoter and
CC terminator in plasmid PLGVneol103 to transform plants, improving
XX nutritional value.
SQ Sequence 1786 BP; 524 A; 452 C; 408 G; 402 T; 0 other;

Query Match 10.9%; Score 201.8; DB 11; Length 1786;
Best Local Similarity 51.6%; Pred. No. 1.8e-39;
Matches 701; Conservative 0; Mismatches 597; Indels 61; Gaps 8;

QY 83 aggagaatgcgtgccagttccagcgccctcaatgscgcagagacctgacaaccgcattgaat 142
Db 132 agtcaacgagtgccaaactcaacaacctcaacgcgttggaaacccgaccacgcgttgagt 191
QY 143 cggaggcggttacaattgagacttggaaacccaacaaccagaggttcgaattgcgcggcg 202
Db 192 ccgaagggtgtcttattgaacatggaactctcaacacctgagctgcaatgscgggtg 251
QY 203 tcgcctctctcgtcttagtctcctccgcgcgcaacgccttcgtlagccttctactccaatg 262
Db 252 tcaactgttccaacagcaccctcaacgcgaacgcgtctccacttgcactcttacttaacct 311
QY 263 ctcccagagagatcttcatccaagcaaggaaggatcttggttgatatctccctggt 322
Db 312 atcccgaatgatcatctctgttcaagggaaggagcaattgtgattgcattccgggat 371
QY 323 gtcttagcacatatgaagagcctgcacaacaaggagccgatatacagttcccaaaagccac 382
Db 372 gtcccgagacglttgagaagccacaacaacaatacaga----- 412
QY 383 caagacgttttgcaagaagaagcaaccaagccaacagcaacaagatgtcaccaagaagtc 442
Db 413 -----agaggtcctaaggttcacagcagcaactacaagacagtcacccaagaagattc 461
QY 443 accgttcaatgagggtgatctcatctgcaagttcccaaccggtgtgtcttcttgcgtgaca 502
Db 462 gtcaacttcaatgaaggagactactagtgatctctcttgggttcttacttgaacctata 521
QY 503 acgaaccaagacactgatgtgtgtgtgttcttcttacttgcacaacaacaacgaacaacc 562
Db 522 acaactgcgatgaaccagttgttgcattgacttcttgcacactccaacttcaacaatc 581
QY 563 agcttgatcagttcccagagagattcaatttgcgtgggaaccagcagaagagttcttaa 622
Db 582 agcttgatcaaaaacccagaglattttaccttgcgtgggaaccagatatagagcatcccg 641
QY 623 gttaccagcaacaagcagacaagcagcagaagaagcttaccatatagcccatacagcc 682
Db 642 agaccatgcaacaa--cagcagcagcagaagagtcattgtgacgcgaagcaggggcaaca 699
QY 683 cgcatagtcgg-cctagacgagaagagcgtgaatttcgcctcgaggaagcagcacagccgc 741
Db 700 ccgacagcaggaggaagaaggtgcaagtgtgtctcagtggttcagcaaaacatttctlagc 759
QY 742 agagaacgagcaggaacaagaagaagaagaggtgnaaatcttcag---cggtc 797
Db 760 acaatcttcaacaccaacgagcagacacagctgagaaacttcggttccagatgacgaag 819

| | | | | |
|----|------|--|--------------|------|
| QY | 798 | caagccgagcttcctctggaacaagccttcacaggttgacgacagacagattgtgc | aaatct | 857 |
| Db | 820 | gaagcagatcgtgacagtgagagggtctcagcgttatcatcagccccaagtygc | aaagaca | 879 |
| QY | 858 | gtgggcgagaaacgagagtgaaagagggagccattgtgacggtgagggg | ccctcag | 917 |
| Db | 880 | agaagacggaagacggaagacggaagaaatatgacgagcccttattcctc | acag | 939 |
| QY | 918 | aatcctgagcccaagatggaacgagaggtgcc--gacgaaagagggatac | gatgaaga | 974 |
| Db | 940 | acgaccaagccatggaagcatgaagatgacgaggaagcgaagaaagaat | caacc | 999 |
| QY | 975 | tcaatatgaataccatgaaacagatggaag-----gcgtggcagggg | | 1016 |
| Db | 1000 | tcgtctgtatcacccctccacagcgacccaaggaagcccgaaacaagaac | caacgltggaag | 1059 |
| QY | 1017 | aagcagagcgcggggggaatggtattgaagagacgattgcaccgcatgtgt | aaagaa | 1076 |
| Db | 1060 | aggaatgcagactagaaatggtgttgaggaataatatlgcaccatgaagct | taacgagaa | 1119 |
| QY | 1077 | cattggtgaaacaagatcccccacacatctacgattccctcagcgctgttc | acccaactg | 1136 |
| Db | 1120 | cattgctgcaccttcacgtgtgacttctacaacccaaaagctggtgc | atlagcaccct | 1179 |
| QY | 1137 | ccacgattcaacctctaatcccttagtgtgttgacctagtgtgaatatg | gaaatct | 1196 |
| Db | 1180 | caacagctcacccctccagccctccgccaattcgactcagtgcccaatat | gtgtct | 1239 |
| QY | 1197 | ctacaggaatgcatgtttgtgccctcactacaacaaacgacacagcatcat | atatgc | 1256 |
| Db | 1240 | ctacaggaatggaatttactctccagattggaacttgaaacgcgaacag | tgtgacg-atga | 1298 |
| QY | 1257 | attgaggggacgggctcaagctgcaagtggtgacagcaacgccaacag | agtgtacagca | 1316 |
| Db | 1299 | ctcgaggggaaaggaaga--gttagagtggtgaaactgccaagggaat | gcagtggtcga | 1356 |
| QY | 1317 | ggagcttcaagaggtcaagttcttgtgtgccaacagaacttcgccgtgtg | ctgggaagtc | 1376 |
| Db | 1357 | tgaagctaaaggagggaacaattgtcagtgtgcccgcagaaaccccg | cggtgtgctga | 1416 |
| QY | 1377 | ccagagcggaacttcgaatacgtggcattccaagacaga | | 1415 |
| Db | 1417 | ggagaaacaagaqattggaaatgtagtggttcaagacaca | | 1455 |

| | | |
|-----------|--|--------------------------|
| RESULT | 13 | |
| AAAN60939 | | |
| ID | AAAN60939 | standard; cDNA; 1786 BP. |
| XX | | |
| AC | AAAN60939; | |
| XX | | |
| DT | 16-OCT-1991 | (first entry) |
| XX | | |
| DE | Sequence encoding soybean glycinin A3B4 subunit. | |
| XX | | |
| KM | Soybean protein; glycinin. | |
| XX | | |
| OS | Glycine max. | |
| XX | | |
| PH | Key | Location/Qualifiers |
| FT | CDS | 47..1597 |
| FT | | /*tag= a |
| XX | | |
| PN | JP61132189-A. | |
| XX | | |
| PD | 19-JUN-1986. | |
| XX | | |
| PE | 03-DEC-1984; | 84JP-0254217. |
| XX | | |
| PR | 03-DEC-1984; | 84JP-0254217. |
| XX | | |
| PA | (NORQ) | NORIINSHO KK |

| | |
|----|--|
| XX | WPI; 1986-200545/31. |
| DR | P-PSDB; AAP61362. |
| XX | |
| PT | Prepn. of soybean messenger RNA - for insertion into cells or |
| PT | microorganisms to produce soybean protein. |
| XX | |
| PS | Example 1; Fig 1; 7pp; Japanese. |
| XX | |
| CC | Sequence derived from mRNA may be used for the expression of the |
| CC | soybean protein by a foreign host. |
| XX | |
| SQ | Sequence 1786 BP; 524 A; 465 C; 395 G; 402 T; 0 other; |

| | | | | |
|---------------------------|--------|--------------------|------------|--------------|
| Query Match | 10.3%; | Score 191.4; | DB 7; | Length 1786; |
| Best Local Similarity | 50.2%; | Pred. No. 6.2e-37; | | |
| Matches 682; Conservative | 0; | Mismatches 616; | Indels 61; | Gaps 6; |

| | | | | | | | | |
|----|-----|-------------------------|---------------------|-----------------|-------------|----------------|------------------------------|-----|
| QY | 83 | aggagaatgctgtccagttccaa | gcgctcaatg | cgagagac | ctgacaac | cgcatlgaat | 142 | |
| | | | | | | | | |
| Db | 132 | agttcaacgagtgccaa | ctcaacaacctcaac | gcgttgyaac | ccgcgaac | ccgcgttgaat | 191 | |
| QY | 143 | cggagggcggttacattgagact | tggaaccccaacaac | ccaggagttcg | aatg | cgccg | 202 | |
| | | | | | | | | |
| Db | 192 | ccgaaggtgtgtcttattg | aaacatgygaactctcaac | accctgagtcgaatg | ccccg | gtg | 251 | |
| QY | 203 | tcgcccctctcgcttagtcc | tcgcgcgcgaac | gcgccttcgt | atagccctt | ctactccaatg | 262 | |
| | | | | | | | | |
| Db | 252 | tcactgtttccaacgcac | ccctcaaccgcgaac | gcgtcc | ccacttgc | catcttacttacc | 311 | |
| QY | 263 | ctcccagagatcttcaatcc | aaggaagggatactt | ggttga | tattcc | ctg | 322 | |
| | | | | | | | | |
| Db | 312 | atcccgaatgatcatgtgtc | gttcaagggaggaatg | gaattg | gat | ttccg | 371 | |
| QY | 323 | gtcctagcaatatgaa | gagcctgcacaaca | aaggaagcgcgat | atcag | tcccaag | 382 | |
| | | | | | | | | |
| Db | 372 | gtcccgaagcgtttg | aagccacaacaac | aatcaagcaga | ----- | | 412 | |
| QY | 383 | caagacgtttgcaaga | aagaccacaagccaac | agcaacaagatag | tccag | aaggtgc | 442 | |
| | | | | | | | | |
| Db | 413 | ----- | agaggtc | aaggtc | acaggtc | acagcaactacaag | agatgc | 461 |
| QY | 443 | accggttcaatgaggtg | tatctcat | tgcagttcc | accggtgtgtc | tttcgtg | taca | 502 |
| | | | | | | | | |
| Db | 462 | gtcaactcaatgaa | ggagagacgtactag | tatccctt | gtgttcc | tactac | ggaacctata | 521 |
| QY | 503 | acgaccacgaacactg | atgtgtgtcgttctct | tactgacac | caacaacga | cgacaac | 562 | |
| | | | | | | | | |
| Db | 522 | acaactg | cgatgaaccagttgt | gtccatcag | tccctcttg | acacctc | caacttcaacaatc | 581 |
| QY | 563 | agcttgatcagttcc | ccaaggaagattc | aat | tgtgc | tgggaacc | agcaaggaagttcttaa | 622 |
| | | | | | | | | |
| Db | 582 | agcttgatcaaaa | cccagagtat | ttac | ctgtc | tyggaac | ccagatatagagcatcccg | 641 |
| QY | 623 | ggtaccagcaaca | aagcagaca | aagcagaagaag | cttacc | ata | tagcccatacagcc | 682 |
| | | | | | | | | |
| Db | 642 | agaccatgcaaca | a-- | cagcagcagcaga | agatcatg | tygacgcaag | cgaggygcaaca | 699 |
| QY | 683 | cgcatagtcgc | gctagacga | aagagcgtga | attcgc | ccctcg | aggaagcaagccgcga | 742 |
| | | | | | | | | |
| Db | 700 | ccgacagcagga | ggaacaag | gtg | catgtgtc | agtyg | cttcaagcaaacattcttagc | 759 |
| QY | 743 | gagaacgagcag | gac----- | aagaaga | agaagacgaag | tygaa | acatctc | 797 |
| | | | | | | | | |
| Db | 760 | acaatccttca | acacccaacga | gagacacag | ctgtg | aga | aacttcgltccagatgacgaag | 819 |
| QY | 798 | cacgcgga | gttccctg | gaacaagc | cttccag | gttg | acgaagacagacagattgtgcaaaatct | 857 |
| | | | | | | | | |
| Db | 820 | gaagcagatc | gtgacag | tgccagg | gagc | ctcagc | gtytatcagcccccaag | 879 |
| QY | 858 | gtggg | gcgagaacgag | aggtg | aaagaagggag | ccattgtg | acggtgagggagcctcag | 917 |

| | | | |
|----|------|---|------|
| Db | 880 | agaaagacgaagacgaagacgaagacgaatatgtgacgcgcctcttatcctccacg | 939 |
| QY | 918 | aatcttgagcccgagatggaacgagaggtgcccgcacgaagaagagatatc---gatgaaga | 974 |
| Db | 940 | acgaccaagccatggaagcatgaagatgacgcgagacggaagccaacaacaagatcaacc | 999 |
| QY | 975 | tcaatatgataaccatcgaacagagatggaagcgctg-----gcagggg | 1016 |
| Db | 1000 | tcgtctgtatcacccctccacagcgaccgaagcaggcccgaaacaagaaccacgtggaag | 1059 |
| QY | 1017 | aagcagagcgcgggggaatgttatgaaagagacatcgcaccgcgatgtgttaaaaaaga | 1076 |
| Db | 1060 | aggaatgcagactagaataatgggttgagaaatatattgcaccatgaagcttcacgagaa | 1119 |
| QY | 1077 | cattggtgaaacagatccccctacacatctacgatacctccagcgctgttcaactcaaatcg | 1136 |
| Db | 1120 | catgtctgcgccttcacgctgtgactcttacaacccaagaagctgtgcgcattagcacccct | 1179 |
| QY | 1137 | ccacgatactcaacctcttaalccttagtgtgtgacttagtgtgaatatggaatct | 1196 |
| Db | 1180 | caacagctctcacccctccagccctccgcgaacttcgactcagtgcccaatatgtgtcct | 1239 |
| QY | 1197 | ctacaggaatgcattgttgtccctcactacatacaaccaacgaacgacacagcatcatatgc | 1256 |
| Db | 1240 | ctacaggaatggaattactctccagatttggaacttgaaagcgacaagtgtagacgatgac | 1299 |
| QY | 1257 | attgaggggagcggtctcacgctgcaagtggtggacagcaacgycacaacagagtgtagcaga | 1316 |
| Db | 1300 | tc---gagccaagaaggaagagttagagtggtgaactgccaaagggaatgcagtgctcgacgg | 1356 |
| QY | 1317 | ggagcttcaagaagggtcacgcttctgtgtggtccacagaacactcgccgtgtggygaagtc | 1376 |
| Db | 1357 | tgaagtaaggaggggacaactgtcagtgtgtccgcagaaaccccgcggtgtgctgagcaag | 1416 |
| QY | 1377 | ccagagcgagaacttcgaatacgtgcatcacaagacaga | 1415 |
| Db | 1417 | gggagacaacaaggaattggaatatgtatgtatgttcaaacacaca | 1455 |

| | |
|-----------|---|
| CC | Glycyltin precursor gene may be used with exogenous promoter and terminator in plasmid pLGVneol103 to transform plants, improving |
| PS | Claim 5; Page 561; 20pp; Japanese. |
| XX | |
| PT | Recombinant plasmid - obtd. by recombining glycyltin gene to |
| PT | plasmid for plant. |
| XX | |
| XX | |
| DR | WPI; 1990-228488/30. |
| XX | |
| PA | (NORO) NORINSHO. |
| PA | (KIRI) KIRIN BREWERY KK. |
| XX | |
| XX | |
| PR | 08-DEC-1988; 88JP-0310553. |
| XX | |
| PF | 08-DEC-1988; 88JP-0310553. |
| XX | |
| PD | 15-JUN-1990. |
| XX | |
| PN | JP02156889-A. |
| XX | |
| OS | Glycine hispida. |
| XX | |
| DE | Glycyltin; glycyltin; pLGVneol103; ds. |
| XX | |
| DT | Glycyltin subunit precursor A5A4B3. |
| XX | |
| DT | 04-DEC-1990 (first entry) |
| XX | |
| AC | AAQ05358; |
| XX | |
| ID | AAQ05358 standard; DNA; 1899 BP. |
| XX | |
| AAQ05358 | |
| RESULT 14 | |

| | |
|----|--|
| CC | nutritional value. |
| XX | |
| SQ | Sequence 1899 BP; 583 A; 468 C; 439 G; 408 T; 0 other. |

| | | | | |
|-----------------------|-----------------|--------------------|------------|--------------|
| Query Match | 9.1%; | Score 168.2; | DB 11; | Length 1899; |
| Best Local Similarity | 57.4%; | Pred. No. 2.9e-31; | | |
| Matches 360; | Conservative 0; | Mismatches 238; | Indels 29; | Gaps 22; |

| | | | |
|--------|---------------|---|-----|
| QY | 10 | ctctcttagccttctctttgtcttcttcttcagtctctggaagctagcagcatctccttc | 69 |
| Db | 43 | cttcactctctctcttcttcttcccttgtctgtactcttctgcagctgcatgtcttgctat | 102 |
| QY | 70 | aggcagcagccggaaggaatgcgtgccagttccagcgccctcaatgcgcagagacctgac | 129 |
| Db | 103 | tagctccagc--aagctcaacgagtgcccaactcaacaacctcaacgcttggaaccgcac | 160 |
| QY | 130 | aaccgcattgaatcggagggcggttacattgagacttggaaaccccaacaacaggaattc | 189 |
| Db | 161 | caccgcgttgagtcggaagtggttgattccaacaacatggaacttccaacacctgagctg | 220 |
| QY | 190 | gaatgcgcgcgcgtccctctctcgtctagtccttcgcgcgcgaacgccccttcgtagccct | 249 |
| Db | 221 | aatgcgcgcggtgtcaactgtttccaactcaacctcaacgcgaatgycctccaactgcga | 280 |
| QY | 250 | ttctactccaatgctcccccaggaagatcttcataccaaggaaggagatactttgggtg | 309 |
| Db | 281 | tcttactcaacctatcccccgatgatcatcatgcgccaagggaagagcacttggaglt | 340 |
| QY | 310 | atatccctggttgccttagcacatatgaagagcctgcacacaagagagcgcgatatcag | 369 |
| Db | 341 | gcaatccaagatgtcctgagacgtttgagagccacaagacaatcaacagagaagc | 400 |
| QY | 370 | tcocaaagaccacaagacglttgcaagaagaagaccaaagccaacagcaacaagatagt | 429 |
| Db | 401 | tc-----aagtcgcagaagcagcagcagctacaggaacagt | 433 |
| QY | 430 | caccagaagtgccaccggttccaatgaggggtgatctcatctcagttcccaaccggtgtgct | 489 |
| Db | 434 | caccagaagatcgtcacttccaatgaaggagagcgtactcgtgattcctcctagtgttctc | 493 |
| QY | 490 | ttctgctgtacaacgcagccagcacactgatgtgtgtgctgttctcttactgaaccaac | 549 |
| Db | 494 | tactggaacctataacactgycgalgaaccaagctgtgtgccatcagctctctgaacacctct | 553 |
| QY | 550 | aacaacgacacacagccttgatcagttcccccaggaagatccaattgctggaacacagag | 609 |
| Db | 554 | aacttcaataacacagcttgatcaaacccctaggtattttaccttgcgtggaaccagat | 613 |
| QY | 610 | caagagttcttaagtgtaaccagcaaca | 636 |
| Db | 614 | atagagtaaccacagagaccatgcaaca | 640 |
| RESULT | 15 | | |
| ID | AA082246 | standard; cDNA; 1646 BP. | |
| AC | AA082246; | | |
| DT | 13-DEC-1990 | (first entry) | |
| DE | | Rice storage protein gene. | |
| XX | | | |
| KW | | Rice storage protein gene; soybean storage protein; glycinin; ss; | |
| XX | | | |
| FH | Key | Location/Qualifiers | |
| FT | CDS | 35..1531 | |
| FT | | /*tag= a | |
| XX | | /product=rice storage protein | |
| PN | JP63071181-A. | | |
| XX | | | |

| | |
|----|---|
| PD | 31-MAR-1988. |
| XX | |
| PF | 16-SEP-1986; 86JP-0217556. |
| XX | |
| PR | 16-SEP-1986; 86JP-0217556. |
| XX | |
| PA | (NORU) NORIINSHO KK. |
| XX | |
| DR | WPI; 1988-128943/19. |
| DR | P-PSDB; AAP82755. |
| XX | |
| PT | New recombinant having messenger RNA of rice storage protein - |
| PT | obtd. in heavier fraction than 18S showing crossing over against |
| PT | messenger RNA of soybean storage protein. |
| XX | |
| PS | Disclosure; ; Japanese. |
| XX | |
| CC | A messenger RNA corresp. to RSP has a base sequence of ca. 1.8 Kb |
| CC | and shows crossing over against the mRNA of soybean storage |
| CC | protein, glycinin, and also those having a partly modified structure. |
| XX | |
| SO | Sequence 1646 BP; 499 A; 374 C; 363 G; 410 T; 0 other; |

Db 1449 ataata 1454

```
Search completed: December 27, 2001, 13:15:28
Job time: 14504 sec
```

| | | | | |
|-----------------------|--------------|--------------------|-----------------|--------------|
| Query Match | 8.6%; | Score 158.8; | DB 9; | Length 1646; |
| Best Local Similarity | 53.5%; | Pred. No. 5.4e-29; | | |
| Matches 356; | Conservative | 0; | Mismatches 307; | Indels 3; |
| | | | | Gaps 1; |

| | | | | | | | | | | | |
|----|------|---------------------|--------------------|-------------------|-----------------------|----------------------------|--------------|--------------|-----------------|-----------|------------|
| QY | 872 | agagtgaagaagaggaagc | aatltgtgacggtgaagg | ggagggcctcagaatc | ttgagccag | 931 | | | | | |
| | | | | | | | | | | | |
| Db | 792 | aaatgacc | aaagagga | aaatgtgtccgtgtcg | aaacacgggtccagttgtgtg | cagccat 851 | | | | | |
| QY | 932 | atggaacgagaggtgtccg | acgaagaagagga | taacgatgaagatca | ataatgaataccatg | 991 | | | | | |
| | | | | | | | | | | | |
| Db | 852 | atgc | atc--- | attgcaggtgacgag | aacaggaacagtgc | aatcaagagagcg | ttatcaag 908 | | | | |
| QY | 992 | aacaggatggaagcg | ctgtgcaggggaagc | agagcggtgggggaatg | gtatgtgaagagacga | 1051 | | | | | |
| | | | | | | | | | | | |
| Db | 909 | aagacacata | tcagcaaa | gtcaatatgtgaag | tgtgtctctaacg | tttgtatgagacct 968 | | | | | |
| QY | 1052 | tctgcaccgc | catgtgtttaaaaa | gaacattgtgtgaa | aacagatcccc | taacatactacgatc 1111 | | | | | |
| | | | | | | | | | | | |
| Db | 969 | tttgacacc | ctgaggtgaaggca | aaacatcgcataatc | ctaaccgtgtgtg | atatacatacaatc 1028 | | | | | |
| QY | 1112 | ctcagcgctgtt | ctcactcaaaa | actgtccacgatctc | aaccttcta | atccctaggtgtg 1171 | | | | | |
| | | | | | | | | | | | |
| Db | 1029 | caagagctgga | aggtttaca | aaatctcaacaccc | agaatttcc | ccattcttagtctgtac 1088 | | | | | |
| QY | 1172 | gacttagtgtga | ataatgaaatc | ctctacaggaatgc | attgttgtcc | ctaataca 1231 | | | | | |
| | | | | | | | | | | | |
| Db | 1089 | agatgag | tgcagtc | aaagtaaatctata | accagaatgc | atcccttcaaccatttgaaca 1148 | | | | | |
| QY | 1232 | ccaacgcacac | agcatcata | atgtgaltgag | ggagcggtc | acgtgcaagtgtgaca 1291 | | | | | |
| | | | | | | | | | | | |
| Db | 1149 | tcaacgctc | acagcgctgtg | tataatactca | aggcggtgccg | gttccaagttgtcaaca 1208 | | | | | |
| QY | 1292 | gcaacgc | caacacagag | tgtacgacgag | cttcaagaggt | lcaacgttctgtgtgccac 1351 | | | | | |
| | | | | | | | | | | | |
| Db | 1209 | acaatgga | aaagacag | tgttcaacg | cgagcttcgcgcg | gacacgtgtgtgtgccac 1268 | | | | | |
| QY | 1352 | agaacttcgc | gtgtgtgtgga | agttcc | cagagcgagaa | cttcgaatacgttgc | atcaaga 1411 | | | | |
| | | | | | | | | | | | |
| Db | 1269 | aacacat | gtcagttgt | aaagaaagcc | aaagagacatgtgt | ctataatgc | atcaaga 1328 | | | | |
| QY | 1412 | cagat | tccaagggcc | acgcatagcca | acttgc | cggtgaa | aaactc | ottcatagataa | actgc 1471 | | |
| | | | | | | | | | | | |
| Db | 1329 | ccaat | cttaactct | atgtgtga | agcca | acattgc | agga | aaagttcc | atcttccgtgt | ctcc 1388 | |
| QY | 1472 | cgagag | gaggtgtgtg | ccaat | ctataatg | gcctcc | caagggag | cgacgca | ggcgttaaga 1531 | | |
| | | | | | | | | | | | |
| Db | 1389 | caactgt | atgttct | tagca | aaatgc | ataatgc | atctc | caagga | agaggtc | agaggtc | aaagc 1448 |
| QY | 1532 | acaaca | 1537 | | | | | | | | |
